

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:04:54 ; Search time 11 Seconds
(without alignments)
27,566 Million cell updates/sec

Title: US-09-824-053-3

Perfect score: 77

Sequence: 1 DLPSPKRVIASNLKF 16

Gaping table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues

Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCYUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
75	97.4	16	10	US-09-824-053-3	Sequence 3, Appl1
75	97.4	546	10	US-09-998-284-2	Sequence 2, Appl1
75	97.4	546	10	US-09-824-053-31	Sequence 31, Appl1
74	96.1	16	10	US-09-824-053-10	Sequence 10, Appl1
5	51.9	1884	10	US-09-785-770A-17	Sequence 17, Appl1
6	40	1907	10	US-09-785-770A-16	Sequence 16, Appl1
7	39	50.6	160	US-09-764-864-1020	Sequence 1020, Ap
8	39	50.6	442	US-09-731-872-286	Sequence 286, App
9	39	50.6	468	US-09-731-872-283	Sequence 283, App
10	39	50.6	468	US-09-933-561-2	Sequence 2, Appl1
11	38	49.4	609	US-09-815-242-10740	Sequence 10740, A
12	38	49.4	793	US-09-925-302-780	Sequence 780, App
13	38	49.4	4545	US-09-873-403-2	Sequence 2, Appl1
14	38	49.4	44	US-09-864-761-36590	Sequence 36590, A
15	37	48.1	119	US-09-867-550-246	Sequence 246, App
16	37	48.1	607	US-09-815-242-13379	Sequence 13379, A
17	37	48.1	607	US-09-815-242-13682	Sequence 13682, A
18	37	48.1	637	US-09-815-242-12058	Sequence 12058, A
19	37	48.1	637	US-09-815-242-12058	Sequence 12058, A

20	36	46.8	388	9	US-09-975-139-9	Sequence 9, Appl1
21	36	46.8	669	10	US-09-801-196-28	Sequence 28, Appl1
22	36	46.8	4636	10	US-09-835-996A-33	Sequence 33, Appl1
23	35.5	46.1	516	9	US-10-067-534-3	Sequence 3, Appl1
24	35.5	46.1	520	9	US-10-067-534-4	Sequence 4, Appl1
25	35	45.5	91	10	US-09-864-761-36754	Sequence 36754, A
26	35	45.5	230	9	US-10-007-814-8	Sequence 8, Appl1
27	35	45.5	420	9	US-10-007-814-6	Sequence 6, Appl1
28	35	45.5	461	9	US-09-738-626-5041	Sequence 5041, Ap
29	35	45.5	503	9	US-10-007-814-2	Sequence 2, Appl1
30	35	45.5	504	9	US-10-007-814-4	Sequence 4, Appl1
31	34.5	44.8	419	10	US-09-741-669-415	Sequence 415, App
32	34	44.2	156	10	US-09-864-761-39604	Sequence 39604, A
33	34	44.2	217	9	US-09-738-626-5837	Sequence 5837, Ap
34	34	44.2	260	9	US-09-764-868-1030	Sequence 1030, Ap
35	34	44.2	451	9	US-09-738-626-3550	Sequence 3550, Ap
36	34	44.2	618	10	US-09-817-676A-14	Sequence 14, Appl
37	34	44.2	618	10	US-09-970-516-4	Sequence 4, Appl1
38	34	44.2	2273	10	US-09-995-542-12	Sequence 12, Appl
39	34	44.2	2310	10	US-09-995-542-10	Sequence 10, Appl
40	34	44.2	26926	9	US-09-759-508B-2	Sequence 2, Appl1
41	33	42.9	15	12	US-10-085-027-4	Sequence 4, Appl1
42	33	42.9	64	10	US-09-864-761-43603	Sequence 43603, A
43	33	42.9	64	10	US-09-764-871-1246	Sequence 1246, Ap
44	33	42.9	104	9	US-09-796-692-1793	Sequence 1793, Ap
45	33	42.9	104	9	US-09-796-692-2057	Sequence 2057, Ap

ALIGNMENTS

RESULT 1
US-09-824-053-3
Sequence 3, Application US/09824053
Patent No. US20020106725A1

GENERAL INFORMATION:
APPLICANT: Peter Stougaard
Ole Cal Hansen

TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME

NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hulton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,053
FILING DATE: 03-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/669,304
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Stanislaus Aksman
REGISTRATION NUMBER: 28,562
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-1926
TELEFAX: (202) 778-2201

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-824-053-3

Query Match 97.4%; Score 75; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DLPMSPRGVASNLKF 16
Db 1 DLPMSPRGVASNLKF 16

RESULT 2
US-09-998-284-2
Sequence 2, Application US/09998284
Patent No. US20020106361A1
GENERAL INFORMATION:
APPLICANT: POULSEN, et al.
TITLE OF INVENTION: COMPOSITION
FILE REFERENCE: 674509-2035
CURRENT APPLICATION NUMBER: US/09/998,284
CURRENT FILING DATE: 2001-11-30
PRIORITY APPLICATION NUMBER: PCT/IB00/00829
PRIORITY FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: GB 9913050.2
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 546
TYPE: PRT
ORGANISM: Chondrus crispus
US-09-998-284-2

Query Match 97.4%; Score 75; DB 10; Length 546;
Best Local Similarity 93.8%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DLPMSPRGVASNLKF 16
Db 219 DLPMSPRGVASNLHF 234

RESULT 3
US-09-824-053-31
Sequence 31, Application US/09824053
Patent No. US20020106725A1
GENERAL INFORMATION:
APPLICANT: Peter Stougaard
Ole Cai Hansen
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hutton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,053
FILING DATE: 03-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/669,304

US-09-824-053-31
Sequence 31, Application US/09824053
Patent No. US20020106725A1
GENERAL INFORMATION:
APPLICANT: Peter Stougaard
Ole Cai Hansen
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hutton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,053
FILING DATE: 03-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/669,304

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Stanislaus Aksman
REGISTRATION NUMBER: 28,562
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-1926
TELEFAX: (202) 778-2201
TELEX: No. US20020106725A1e
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 546 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-824-053-31

Query Match 97.4%; Score 75; DB 10; Length 546;
Best Local Similarity 93.8%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DLPMSPRGVASNLKF 16
Db 219 DLPMSPRGVASNLHF 234

RESULT 4
US-09-824-053-10
Sequence 10, Application US/09824053
Patent No. US20020106725A1
GENERAL INFORMATION:
APPLICANT: Peter Stougaard
Ole Cai Hansen
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hutton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,053
FILING DATE: 03-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/669,304
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Stanislaus Aksman
REGISTRATION NUMBER: 28,562
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-1926
TELEFAX: (202) 778-2201
TELEX: No. US20020106725A1e
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-824-053-10

US-09-824-053-10
Sequence 10, Application US/09824053
Patent No. US20020106725A1
GENERAL INFORMATION:
APPLICANT: Peter Stougaard
Ole Cai Hansen
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hutton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/824,053
FILING DATE: 03-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/669,304
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Stanislaus Aksman
REGISTRATION NUMBER: 28,562
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-1926
TELEFAX: (202) 778-2201
TELEX: No. US20020106725A1e
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-824-053-10

Query Match 96.1%; Score 74; DB 10; Length 16;
Best Local Similarity 93.8%; Pred. No. 7.4e-07;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DLPMSRGVYASNLXF 16
|||||

Db 1 DLPMSRGVYASNLWF 16

RESULT 5
US-09-785-770A-17
; Sequence 17, Application US/09785770A
; Patent No. US20020103360A1
; GENERAL INFORMATION:
; APPLICANT: Barnes, Thomas M.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: A NOVEL PROTEIN RELATED TO MELANOMA-INHIBITING PROTEIN
; FILE REFERENCE: 07334-328001
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/387,462
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/145,056
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 17
; LENGTH: 1884
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-770A-17

Query Match 51.9%; Score 40; DB 10; Length 1884;
Best Local Similarity 43.8%; Pred. No. 90;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 DLPMSRGVYASNLXF 16
|||||

Db 1804 DLPMSRGVYASNLWF 1819

RESULT 6
US-09-785-770A-16
; Sequence 16, Application US/09785770A
; Patent No. US20020103360A1
; GENERAL INFORMATION:
; APPLICANT: Barnes, Thomas M.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: A NOVEL PROTEIN RELATED TO MELANOMA-INHIBITING PROTEIN
; FILE REFERENCE: 07334-328001
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/387,462
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/145,056
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 16
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-770A-16

Query Match 51.9%; Score 40; DB 10; Length 1907;
Best Local Similarity 43.8%; Pred. No. 92;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 DLPMSRGVYASNLXF 16
|||||

Db 1827 DLPMSRGVYASNLWF 1842

RESULT 7
US-09-764-864-1020
; Sequence 1020, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1020
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1020

Query Match 50.6%; Score 39; DB 10; Length 160;
Best Local Similarity 70.0%; Pred. No. 8.8;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 DLPMSRGVYASNL 11
|||

Db 137 DLPMSRGVYASNL 146

RESULT 8
US-09-731-872-286
; Sequence 286, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouguetere, Lydie
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 286
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -21..-1
; NAME/KEY: UNSURE
; LOCATION: 132
; OTHER INFORMATION: Xaa = Pro,Arg
US-09-731-872-286

Query Match 50.6%; Score 39; DB 10; Length 442;
Best Local Similarity 37.5%; Pred. No. 27;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 DLPMSRGVYASNLXF 16
|||

Db 176 DLPMSRGVYASNLWF 191

RESULT 9
US-09-731-872-283
; Sequence 283, Application US/09731872

Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Joberet, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.053.REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 283
LENGTH: 468
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -21...-1
9-731-872-283

Query Match 50.6%; Score 39; DB 10; Length 468;
Best Local Similarity 37.5%; Pred. No. 29;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 DLPMSPRGVASNLXF 16
DB 176 NLPLSPQGTVRTAVEF 191

RESULT 10
US-09-933-561-2
Sequence 2, Application US/09933561
Patent No. US2002010664A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
APPLICANT: Kaser, Matthew R.
APPLICANT: Baugh, Mariah R.
TITLE OF INVENTION: TAPASIN-LIKE PROTEIN
FILE REFERENCE: PC-0002-1 CIP
CURRENT APPLICATION NUMBER: US/09/933,561
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 09/292,097
PRIOR FILING DATE: 1999-04-14
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 468
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 103348CD1
09-933-561-2

Query Match 50.6%; Score 39; DB 10; Length 468;
Best Local Similarity 37.5%; Pred. No. 29;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 DLPMSPRGVASNLXF 16
DB 176 NLPLSPQGTVRTAVEF 191

RESULT 11
US-10-090-185-10
Sequence 10, Application US/10090185
Publication No. US20020197647A1
GENERAL INFORMATION:
APPLICANT: Zhang, Xiaokui
APPLICANT: Wieszczynska, Melissa H

APPLICANT: Horvath, Curt M
APPLICANT: Darnell Jr., James E
TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVITY
TITLE OF INVENTION: INTERACTIONS
FILE REFERENCE: 600-1-253
CURRENT APPLICATION NUMBER: US/10/090,185
CURRENT FILING DATE: 2002-03-04
PRIOR APPLICATION NUMBER: 09/387,418
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10
LENGTH: 393
TYPE: PRT
ORGANISM: Mus musculus
US-10-090-185-10

Query Match 49.4%; Score 38; DB 9; Length 393;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLPMSPR 7
DB 346 DLPMSPR 352

RESULT 12
US-09-815-242-10740
Sequence 10740, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes In
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10740
LENGTH: 609
TYPE: PRT
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(609)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-10740

Query Match 49.4%; Score 38; DB 10; Length 609;
Best Local Similarity 43.8%; Pred. No. 57;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 DLPMSRGVIASTNMF 16
Db 432 DIPAPRGVQJEVSF 447

RESULT 13
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-780

Query Match 49.4%; Score 38; DB 10; Length 793;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLPMSR 7
Db 746 DLPMSR 752

RESULT 14
US-09-873-403-2
; Sequence 2, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
; APPLICANT: Strivastava, Pramod K
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
; FILE REFERENCE: 8449-178
; CURRENT APPLICATION NUMBER: US/09/873,403
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/625,139
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/209,266
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-403-2

Query Match 49.4%; Score 38; DB 10; Length 4545;
Best Local Similarity 54.3%; Pred. No. 5,2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 DLPMSRGVIA 11
Db 2460 DIPQPMGITA 2470

RESULT 15
US-09-864-761-36590
; Sequence 36590, Application US/09864761
; Patent No. US20020048763A1

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36590
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: MAP TO AC006367.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P77672, EVALUATE 2.10e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE960561.1, EVALUATE 1.80e+00
US-09-864-761-36590

Query Match 48.1%; Score 37; DB 10; Length 44;
Best Local Similarity 58.3%; Pred. No. 4.6;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 PMSPRGVIASTNL 14

Db 4 PMPRGVAVSL 15

RESULT 16

US-09-867-550-246
; Sequence 246, Application US/09867550
; Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie

TITLE OF INVENTION: Topper, James

FILE REFERENCE: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

CURRENT FILING DATE: 2001-09-20

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 246

LENGTH: 119

TYPE: PRT

ORGANISM: Homo sapiens

NAME/KEY: VARIANT

LOCATION: (1)

OTHER INFORMATION: wherein Xaa may be any one of Arg or Gln or Leu or Pro or Ser or

OTHER INFORMATION: Thr or Lys or Ala or Val or Glu or Gly

US-09-867-550-246

Query Match 48.1%; Score 37; DB 10; Length 119;

Best Local Similarity 45.5%; Pred. No. 14;

Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 51 MPRGIVAAV 61

RESULT 17

US-09-815-242-13379
; Sequence 13379, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT FILING DATE: 2001-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13379
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13379

Query Match

Best Local Similarity 48.1%; Score 37; DB 10; Length 607;

Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 431 DIPAPRGIPQIEVTF 446

RESULT 18

US-09-815-242-13682
; Sequence 13682, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT FILING DATE: 2001-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13682

LENGTH: 607

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

US-09-815-242-13682

Query Match

Best Local Similarity 48.1%; Score 37; DB 10; Length 607;

Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 431 DIPAPRGIPQIEVTF 446

RESULT 19

US-09-815-242-12058
; Sequence 12058, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12038
 LENGTH: 637
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-815-242-12058

Query Match 48.18; Score 37; DB 10; Length 637;
 Best Local Similarity 43.88; Pred. No. 89;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 DLPMPRGVIAASNLXF 16
 Db 461 DIPAPRGVPOIEYTF 476

RESULT 20
 US-09-975-139-9
 Sequence 9, Application US/09975139
 Patent No. US20020155460A1
 GENERAL INFORMATION:
 APPLICANT: Genencor International, Inc.
 APPLICANT: Schellenberger, Volker
 APPLICANT: Nakl, Donald
 APPLICANT: Morrison, Thomas B.
 TITLE OF INVENTION: INFORMATION RICH LIBRARIES
 FILE REFERENCE: 23623-7060
 CURRENT APPLICATION NUMBER: US/09/975,139
 CURRENT FILING DATE: 2001-10-10
 PRIOR APPLICATION NUMBER: US 60/239,476
 PRIOR FILING DATE: 2000-10-10
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 388
 TYPE: PRT
 ORGANISM: Yersinia enterocolitica
 FEATURE:
 OTHER INFORMATION: AmpC protein
 US-09-975-139-9

Query Match 46.88; Score 36; DB 9; Length 388;
 Best Local Similarity 54.58; Pred. No. 76;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 PMSPRGVIAASN 13
 Db 304 PINPQGVIAADS 314

Search completed: January 2, 2003, 12:10:23
 Job time : 12 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:03:01 ; Search time 35 Seconds

(without alignments)
60.915 Million cell updates/sec

Title: US-09-824-053-3

Perfect score: 77

Sequence: 1 DLPSPRGVIAINLXF 16

ling table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A.Geneseq.101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	97.4	16	AAW20070	HOX3, a hexose oxi
2	75	97.4	18	AAW20076	Hexose oxidase, an
3	75	97.4	21	AAW83619	Synthetic hexose o
4	75	97.4	22	AAU02192	Synthetic hexose o
5	75	97.4	22	AAW59205	Chondrus crispus he
6	44	57.1	187	ABP30454	Streptococcus poly
7	44	57.1	230	AAV91319	Group B Streptococ
8	44	57.1	230	ABP28247	Streptococcus poly
9	42	54.5	2506	ABG07191	Novel human diagno
10	42	54.5	22	ABG30064	Novel human diagno

11	40	51.9	185	22	ABG23334	Novel human diagno
12	40	51.9	245	23	ABP41781	Human ovarian anti
13	40	51.9	255	23	AAU72763	Tomato DM7 protein
14	40	51.9	1193	22	AAW25602	Human protein sequ
15	40	51.9	1193	23	ABG61824	Prostate cancer as
16	40	51.9	1194	22	AAU32407	Novel human secret
17	39	50.6	70	22	ABG24042	Novel human diagno
18	39	50.6	84	22	AAU22735	Human prostate can
19	39	50.6	97	22	AAW94707	Human reproductive
20	39	50.6	103	22	AAU20801	Human immune/haema
21	39	50.6	103	22	AAU20801	Human novel foetal
22	39	50.6	113	22	AAU23057	Novel human enzyme
23	39	50.6	143	22	ABB12095	Human secreted pro
24	39	50.6	144	8	AAU70471	Sequence encoded b
25	39	50.6	148	8	AAU70471	Human testicular a
26	39	50.6	149	22	ABW65590	Six repeats of seq
27	39	50.6	149	22	AAW96591	Human secreted pro
28	39	50.6	151	22	ABG26574	Novel human diagno
29	39	50.6	152	22	AAW84814	Human immune/haema
30	39	50.6	155	21	AAW35596	Arabidopsis thalia
31	39	50.6	160	22	AAU16067	Human novel secret
32	39	50.6	268	22	AAU20491	Human secreted pro
33	39	50.6	277	22	AAE09683	Human gene 1 encod
34	39	50.6	307	23	AAW90285	Human polypeptide
35	39	50.6	308	20	AAV30811	Human secreted pro
36	39	50.6	319	21	AAW41889	Human protein sequ
37	39	50.6	338	22	AAW40079	Human tapasin-like
38	39	50.6	442	22	AAW89166	Human placenta acu
39	39	50.6	468	22	AAW89166	Human signal trans
40	39	50.6	468	22	AAW89166	Human secreted pro
41	39	50.6	468	22	AAW89166	Human secreted pro
42	39	50.6	468	22	AAW89166	Human secreted pro
43	39	50.6	770	16	AAW82993	Human placenta acu
44	39	50.6	770	23	AAW15174	Human signal trans
45	39	50.6	1132	17	AAW97866	Chicken leucocytos

ALIGNMENTS

RESULT 1	AAW20070	standard; peptide; 16 AA.
ID	AAW20070	
AC	AAW20070	
XX		
DT	12-SEP-1997	(first entry)
XX		
DE	HOX3, a hexose oxidase derived antimicrobial/antioxidant peptide.	
XX		
KW	Hexose oxidase; Chondrus crispus; marine algae; recombinant;	
KW	antimicrobial; antioxidant; food preparation; dairy product; starch;	
KW	beverage; animal feed; silage; sugar reduction; cosmetics; dental;	
KW	toothpaste; dough; lactone production.	
XX		
OS	Chondrus crispus.	
XX		
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 15	/note="x- any naturally occurring amino acid"
XX		
PN	WO9640935-A1.	
XX		
PD	19-DEC-1996.	
XX		
PF	04-JUN-1996;	96WO-DK00238.
XX		
PR	07-JUN-1995;	95US-0476910.
XX		
PA	(BIOT-) BIOTEKNOLOGISK INST.	
XX		
PI	Hansen OC, Stougaard P;	
XX		

DR WPI; 1997-052332/05.
XX Recombinant production of polypeptide having hexose oxidase activity
PT - used in food preparations as antibacterial and antioxidant agent
XX
PS Claim 6; Page 99; 124pp; English.
XX
CC AAM20068-75 are small peptides derived from hexose oxidase (HO) of the
CC marine algae species Chondrus crispus. Hexose oxidase and peptides
CC having HO activity are useful in the production of food products, e.g.
CC dairy products, starch-containing food products (dough) and non-dairy
CC beverages. HO and active peptide fragments have antimicrobial and
CC antioxidant properties and act by removing all the oxygen in a food
CC packaging. HO and peptides of HO can also be used in an animal feed,
CC especially silage. Further uses are to reduce, or analyse, the sugar
CC content in a food, in the production of cosmetics, tooth care products
CC or a pharmaceutical product and in lactone production. HO can be
CC recombinantly produced in industrially appropriate quantities, and at
CC a quality and purity level which renders polypeptide suitable for
CC industrial purposes.

Sequence 16 AA;
Query Match 97.4%; Score 75; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPRGVASNLXF 16
Db 1 DLPMSPRGVASNLXF 16

RESULT 2
AAM20076
ID AAM20076 standard; Protein: 546 AA.
XX
AC AAM20076;
XX
DT 12-SEP-1997 (first entry)
XX
DE Hexose oxidase, an antimicrobial/antioxidant agent.
XX
XX Hexose oxidase; Chondrus crispus; marine algae; recombinant;
XX antimicrobial; antioxidant; food preparation; dairy product; starch;
XX beverage; animal feed; silage; sugar reduction; cosmetics; dental;
XX toothpaste; dough; lactone production.
XX
OS Chondrus crispus.
XX
PN WO9640935-A1.
XX
XX 19-DEC-1996.
XX
XX 04-JUN-1996; 96WO-DK00238.
XX
XX 07-JUN-1995; 95US-0476910.
XX
XX (BIOT-) BIOTEKNOLOGISK INST.
XX
XX Hansen OC, Stougaard P;
XX
XX WPI; 1997-052332/05.
XX
XX N-PSDB; AAT76552.
XX
PT Recombinant production of polypeptide having hexose oxidase activity
PT - used in food preparations as antibacterial and antioxidant agent
XX
XX Example 3.5; Page 95-97; 124pp; English.
XX
XX AAM20076 shows the hexose oxidase (HO) enzyme of the marine algae
XX species Chondrus crispus. HO is useful in the production of food
XX products, e.g. dairy products, starch-containing food products (dough)
XX and non-dairy beverages. HO and active peptide fragments have

CC antimicrobial and antioxidant properties and act by removing all the
CC oxygen in a food packaging. HO and peptides of HO can also be used in an
CC animal feed, especially silage. Further uses are to reduce, or analyse,
CC the sugar content in a food, in the production of cosmetics, tooth care
CC products or a pharmaceutical product and in lactone production. HO can
CC be recombinantly produced in industrially appropriate quantities, and at
CC a quality and purity level which renders the polypeptide suitable for
CC industrial purposes.

Sequence 546 AA;
Query Match 97.4%; Score 75; DB 18; Length 546;
Best Local Similarity 93.8%; Pred. No. 5.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLPMSPRGVASNLXF 16
Db 219 DLPMSPRGVASNLHF 234

RESULT 3
AAY83619
ID AAY83619 standard; Protein: 546 AA.
XX
AC AAY83619;
XX
DT 29-AUG-2000 (first entry)
XX
DE Synthetic hexose oxidase.
XX
XX Hexose oxidase; production; fermentation; synthetic; modification;
XX prokaryote; eukaryote.
XX
OS Synthetic.
XX
XX EP1008651-A2.
XX
XX 14-JUN-2000.
XX
XX 01-DEC-1999; 99EP-0204068.
XX
XX 09-DEC-1998; 98DK-0001630.
XX
XX (BIOT-) BIOTEKNOLOGISK INST.
XX
XX Stougaard P, Pedersen LH, Wolff AM, Poulsen U, Hansen OC;
XX
XX WPI; 2000-389309/34.
XX
XX N-PSDB; AA294011.
XX
XX Nucleic acid fragment useful for producing large amounts of hexose
XX oxidase comprises nucleotide sequence encoding hexose oxidase that is
XX modified by at least one codon
XX
XX Disclosure; Page 28-29; 42pp; English.
XX
XX Wild type hexose oxidase can be produced by isolating a nucleotide
XX sequence coding for hexose oxidase naturally produced by an organism
XX and modifying the sequence so that the modified sequence is
XX expressed, under identical conditions, at a level at least 10% higher
XX than the non-modified sequence. The method is useful for producing
XX hexose oxidase in prokaryotic and eukaryotic cells. Previous methods
XX of producing hexose oxidase were to isolate the enzyme from a source
XX that naturally produces the enzyme, using host organisms which
XX produce a relatively low level of expression. Therefore, industrial
XX production of the enzyme is not feasible using this method. This is a
XX much improved method which produces at least 250 mg enzyme per liter
XX of fermentation medium.
XX
XX Sequence 546 AA;
XX
XX Query Match 97.4%; Score 75; DB 21; Length 546;
XX Best Local Similarity 93.8%; Pred. No. 5.3e-05;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLPMSRGTASNLXF 16
|||||
DB 219 DLPMSRGTASNLHF 234

RESULT 4

AAU02192
ID AAU02192 standard; Protein; 546 AA.

AC AAU02192;

DT 12-SEP-2001 (first entry)

Synthetic hexose oxidase (HOX) amino acid sequence.

HOX; seaweed; hexose oxidase; interleukin I receptor antagonist; IL-1ra;
food manufacturing; beverage; detergent; baking; dough improving agent;
D-hexose:O2-oxidoreductase; EC 1.1.3.5.

OS Chondrus crispus.
Synthetic.

PN WO200138544-A1.

PD 31-MAY-2001.

PF 24-NOV-2000; 2000WO-IB01886.

PR 24-NOV-1999; 99GB-0027801.

PA (DANI-) DANISCO AS.

PI Johansen CL, Kjaerulff S, Madrid SM, Pedersen H, Poulsen CH;

PT Zarqah MR;

DR WPI: 2001-367695/38.
N-PSDB; MAS06173.

XX Releasing soluble or membrane associated intracellular protein from a
PT cell for manufacturing food, comprises contacting the cell with a
PT membrane extracting composition and causing the protein to be released
PT in soluble form -

PS Claim 16; Fig 6; 108bp; English.

CC The sequence represents the amino acid sequence of synthetic hexose
CC oxidase (D-hexose:O2-oxidoreductase, EC 1.1.3.5), also referred to as HOX.
CC The native HOX gene was altered using site-directed mutagenesis in order
CC to match the codon usage to known codon preferences of biotechnologically
CC relevant yeasts, such as *Pichia sp.*, to facilitate high level production
CC of HOX in these organisms. The invention involves a method for releasing
CC a soluble or membrane associated intracellular protein of interest (POI)
CC from a cell involving contacting a cell comprising a soluble or membrane
CC associated intracellular POI with a membrane extracting composition (1)
CC and causing the POI to be released from the cell in a soluble form. The
CC method is useful for releasing POI, such as an interleukin I receptor
CC antagonist (IL-1ra) which involves contacting a transformed cell
CC comprising IL-1ra with (1) and causing IL-1ra to be released from the
CC transformed cell, in a soluble form. The method is also useful for
CC screening mutated cells or transformed cells producing elevated levels of
CC intracellular POI. The method is used to release a POI for manufacturing
CC food products, such as beverages, preparation of detergents, and in
CC baking as a dough improving agent. The method obtains a fast, specific
CC and economically efficient extraction of a soluble or membrane associated
CC intracellular POI without the use of conventional cell disruption
CC techniques. The resulting cell extract contains less contaminating
CC intracellular DNA and is relatively free of cell wall fragments. The
CC intracellular POI can be recovered from a eukaryotic host organism such
CC as yeast, before glycosylation takes place. The method can be used to
CC prevent contact of intracellular POI with the extracellular growth
CC medium.

XX Sequence 546 AA;

QY 1 DLPMSRGTASNLXF 16
|||||
DB 219 DLPMSRGTASNLHF 234

Query Match 97.4%; Score 75; DB 22; Length 546;
Best Local Similarity 93.8%; Pred. No. 5.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5

AAU02192
ID AAU02192 standard; Protein; 546 AA.

AC AAU02192;

DT 23-MAR-2001 (first entry)

DE Chondrus crispus hexose oxidase enzyme protein.

XX Hexose oxidase; marine alga; anti-fouling.

OS Chondrus crispus.

PN WO200075293-A2.

PD 14-DEC-2000.

PF 02-JUN-2000; 2000WO-IB00829.

PR 04-JUN-1999; 99GB-0013050.

PA (DANI-) DANISCO AS.

PI Poulsen CH, Kragh KM;

DR WPI: 2001-112148/12.

XX New anti-fouling composition, useful as a coating for treating
PT different surfaces, e.g. outdoor woodwork, external surface of a
PT central heating system, or a hull of a marine vessel -

PS Claim 5; Page 35-36; 36bp; English.

CC The present invention relates to a new anti-fouling composition. The
CC composition involves a surface coating material, a hexose oxidase
CC enzyme obtained from a marine organism and a substrate for the
CC enzyme. The anti-fouling composition is useful as a coating formulated
CC for treating a surface, e.g. outdoor wood work, external surface of a
CC central heating system, or a hull of a marine vessel. It is
CC also useful as an anti-fouling agent for marine structures exposed to
CC seawater flora and fauna.

QY Sequence 546 AA;

Query Match 97.4%; Score 75; DB 22; Length 546;
Best Local Similarity 93.8%; Pred. No. 5.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLPMSRGTASNLXF 16
|||||
DB 219 DLPMSRGTASNLHF 234

RESULT 6

AAU02192
ID AAU02192 standard; Protein; 187 AA.

AC AAU02192;

DT 02-JUL-2002 (first entry)

XX	Streptococcus	polypeptide SEQ ID NO 10084.
DE		
XX	Streptococcus	GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW	group A streptococcus; Streptococcus pyogenes; antibacterial;	
KM	antiinflammatory; infection; vaccine; meningitis; gene therapy.	
XX		
OS	Streptococcus	agalactiae.
XX		
PN	WO200234771-A2.	
XX		
PD	02-MAY-2002.	
XX		
PE	29-OCT-2001; 2001WO-GB04789.	
XX		
PR	27-OCT-2000; 2000GB-0026333.	
PR	24-NOV-2000; 2000GB-0028727.	
PR	07-MAR-2001; 2001GB-0005640.	
XX		
PA	(CHIR-) CHIRON SPA.	
XX	(GENO-) INST GENOMIC RES.	
XX		
XX	Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;	
XX	Teetelin H;	
DR	WPI; 2002-352536/38.	
DR	N-PSDB; ABN71085.	
XX		
PT	New Streptococcus protein for the treatment or prevention of infection	
PT	or disease caused by Streptococcus bacteria, such as meningitis, and	
PT	for detecting a compound that binds to the protein -	
XX		
PS	Claim 1; Page 4137; 4525pp; English.	
XX		
CC	The invention relates to a protein (ABP25413-ABP30895) from group B	
CC	streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS	
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1) given in	
CC	the specification. The proteins have antibacterial and antiinflammatory	
CC	activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and	
CC	antibodies that bind (I) are used in the manufacture of medicaments for	
CC	the treatment or prevention of infection or disease caused by	
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.	
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a	
CC	biological sample. (I) is used to determine whether a compound binds to	
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be	
CC	used as a vaccine or diagnostic composition. The disease caused by	
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic	
CC	acid encoding (I) may be used to recombinantly produce (I) and may be	
CC	used in gene therapy. Antibodies to (I) are used for affinity	
CC	chromatography, immunoassays, and distinguishing/identifying	
CC	Streptococcus proteins.	
XX		
XX	Sequence 187 AA:	
XX		
XX	ery Match 57.1%; Score 44; DB 23; Length 187;	
XX	est Local Similarity 75.0%; Pred. NO. 6.1;	
XX	Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
OY	5 SPRGVIASNLAF 16	
DB	:	
	13 SPRGVIANKLWF 24	
	:	
RESULT 7		
AAV91319		
ID	AAV91319 standard; Protein; 230 AA.	
XX		
XX	AAV91319;	
XX		
DT	30-MAY-2000 (first entry)	
XX		
DE	Group B Streptococcus protein sequence SEQ ID NO:47.	
XX		

KW	Group B Streptococcus; Streptococcus agalactiae; protein antigen;
RW	vaccine; screening; immunogen; detection; diagnosis; infection;
KM	antibody; affibody; antibacterial.
XX	
OS	Streptococcus agalactiae.
XX	
PN	WO200006736-A2.
PD	
10-FEB-2000.	
XX	
PE	27-JUL-1999; 99WO-GB02444.
XX	
PR	27-JUL-1998; 98GB-0016335.
PR	19-MAR-1999; 99US-0125163.
XX	
PA	(MICR-) MICROBIAL TECHNICS LTD.
XX	
PI	Le Page RWF, Wells JM, Hanniffy SB;
DR	WPI; 2000-195299/17.
PT	New Group B Streptococcus protein, useful as vaccine, for diagnosis of
ST	Streptococcal infections and for screening of antibodies or affibodies
XX	
PS	Claim 1; Fig 1; 123pp; English.
CC	AAA05803 to AAA05872 encode proteins, polypeptides and peptides (given
CC	in AAY1275 to AAY91343) isolated from Group B Streptococcus (GBS), also
CC	known as Streptococcus agalactiae. The GBS polynucleotides and
CC	polypeptides have antibacterial activity. Immunogenic compositions
CC	comprising GBS polynucleotides or polypeptides can be used as vaccines
CC	and for the treatment or prophylaxis of GBS infection. The
CC	polynucleotides and polypeptides can also be used in the detection of GBS
CC	and for screening DNA encoding bacterial cell envelope associated or
CC	secreted antigens in gram positive bacteria. AAA05873 to AAA05941
CC	represent primers used in the exemplification of the present invention.
SO	
Sequence	230 AA;
Query Match	57.1%; Score 44; DB 21; Length 230;
Best Local Similarity	75.0%; Pred. No. 7.7;
Matches	9; Conservative 1; Mismatches 2; Indels 0; Gaps 0.
OY	5 SPRGVASNLXF 16 :
DB	56 SPRGVIANKLVE 67
RESULT 8	
ID	ABP28247 standard; Protein; 230 AA.
AC	ABP28247;
DT	02-JUL-2002 (first entry)
DE	Streptococcus polypeptide SEQ ID NO 5670.
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
RW	group A streptococcus; Streptococcus pyogenes; antibacterial;
KM	antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX	
OS	Streptococcus agalactiae.
XX	
PN	WO200234771-A2.
PD	
02-MAY-2002.	
XX	
PF	29-OCT-2001; 2001WO-GB04789.
PR	27-OCT-2000; 2000GB-0026333.
PR	24-NOV-2000; 2000GB-0028727.
PR	07-MAR-2001; 2001GB-0005640.

XX (CHTR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN68878.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX

Claim 1; Page 3729; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B Streptococcus (Streptococcus agalactiae) or group A Streptococcus (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

Sequence 230 AA;

Query Match 57.1%; Score 44; DB 23; Length 230;

Best Local Similarity 75.0%; Pred. No. 7.7;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 SPRGVIANSLXF 16
 |||||: ||
 Db 56 SPRGVIANSLVF 67

RESULT 9

ABG07191

ABG07191 standard; Protein: 2506 AA.

ABG07191;

13-FEB-2002 (first entry)

Novel human diagnostic protein #7182.

Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

MO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

XX

DR WPI: 2001-639362/73.
 DR N-PSDB; AAS71378.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX

Claim 20; SEQ ID No 37550; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

Sequence 2506 AA;

Query Match 54.5%; Score 42; DB 22; Length 2506;

Best Local Similarity 53.3%; Pred. No. 2.6e+02;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 LPMSPRGVIANSLXF 16
 |||||: ||
 Db 745 LPAPRGVQRLRLF 759

RESULT 10

ABG30064

ABG30064 standard; Protein: 2506 AA.

ABG30064;

18-FEB-2002 (first entry)

Novel human diagnostic protein #30055.

Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

MO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

XX

WPI: 2001-639362/73.

N-PSDB; AAS94251.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20: SEQ ID No 60423: 103bp: English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences).

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy technique to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pat/published_pat_sequences](http://wipo.int/pat/published_pat_sequences).

XX WPI: 2002-147878/19.
DR N-PSDB: ABQ54858.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

PS Claim 11: SEQ ID No 2913; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovarian and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 245 AA:

Query Match 51.9%; Score 40; DB 23; Length 245;
Best Local Similarity 43.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

1 DLPSPRGVIASNLXF 16
||| ||| : :
165 DLPHPRGFLGHAHF 180

RESULT 13

AAU72763 standard; Protein: 255 AA.

AAU72763:

26-FEB-2002 (first entry)

Tomato DMT protein related sequence #2.

XX Demeter: DMT; Atropos: ATR; 5-methylcytosine glycosylase;
XX DNA demethylation; transgenic plant; transcription modulation;
XX flowering time; endosperm development; MEDEA.

XX Lycopersicon esculentum.

XX MO200180626-A1.

XX 01-NOV-2001.

XX 23-APR-2001; 2001MO-US13059.

XX 21-APR-2000; 2000US-0553690.

XX (REGC) UNIV CALIFORNIA.

XX Fischer RL, Choi Y, Hannon M, Okumuro JK, Tatarnova TV;

XX WPI: 2002-055307/07.

XX N-PSDB: AAS96713.

XX New polynucleotide that control plant development comprising a sequence
XX having a specific homology to DEMETER domains A,B or C -
XX Disclosure: Page 96; 109pp; English.

XX The invention relates to an isolated polynucleotide sequence or their
XX complement encoding a polypeptide having a sequence at least 40%
XX identical to DMT (DEMETER, previously known as ATROPOS (ATR)) Domain A, B
XX or C or their combinations. Also included are an expression
XX cassette comprising the polynucleotide or comprising a heterologous
XX polynucleotide under the control of a promoter at least 70% identical to
XX DMT 5' flanking sequence, DMT 3' flanking sequence or an 5' untranslated
XX region of DMT, a host cell comprising an exogenous polynucleotide
XX encoding a DMT-like protein and a transgenic plant comprising a
XX polynucleotide encoding a DMT-like protein. The expression cassette is
XX useful for modulating transcription. The method comprises introducing
XX the cassette into a host cell preferably Agrobacterium by sexual
XX cross, and selecting a host cell with modulated transcription, where
XX the protein is capable of exhibiting at least one of the following
XX biological activities, which include enhanced expression of the
XX protein in a plant results in a delay in flowering time, introduction of
XX the protein into a cell results in modulation of methylation of
XX chromosomal DNA in the cell, reduction of expression of the protein in a
XX plant results in enhanced endosperm development and expressing of the
XX protein in an Arabidopsis leaf results in expression of the MEDEA gene.
XX The polynucleotide is useful for detecting a nucleic acid in a sample.
XX DEMETER is related to 5-methylcytosine glycosylases and regulates
XX transcription of target genes by demethylation. The present
XX sequence represents a DMT-like protein.

XX Sequence 255 AA:

Query Match 51.9%; Score 40; DB 23; Length 255;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DLPSPRGVIASNL 14
||| ||| : :
Db 120 DLPSPSSVSSST 133

RESULT 14

AAAM25602 standard; Protein: 1193 AA.

AAAM25602:

16-OCT-2001 (first entry)

Human protein sequence SEQ ID NO:1117.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX anti-inflammatory; antirheumatic; antiallergic; immunosuppressive;
XX antibacterial; endocrine; cardiant; central nervous system; viticude;
XX anti-HIV; fungicide; antimutagen; cardiovascular; anlaemic; anaemia;
XX antagregant; haemostatic; vulnary; antileuk; osteopathic; eczema;
XX dermatological; antiallergic; antiasthmatic; antidiabetic; cystostatic;
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX thrombocytopenia; osteoporosis; severe combined immunodeficiency;

KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 XX neurological disorder.
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US35017.
 XX
 PR 23-DEC-1999; 99US-0471275.
 XX 21-JAN-2000; 2000US-048725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 PI WPI: 2001-457603/49.
 XX N-PSDB; AAH9543.
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 PS Claim 20; Page 230; 1217pp; English.
 XX
 CC AAH95166 to AAH99904 encode the human proteins given in AAH25225 to
 CC AAH25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antineuritic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC anticancer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytosolic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 CC
 SQ Sequence 1193 AA;
 XX
 QY 1 DLPMSPRGVIASNLF 16
 |||: |||: : :
 ID 1113 DLPMPRGFLPGHAF 1128
 DB
 RESULT 15
 ABG61824
 ID ABG61824 standard; Protein: 1193 AA.
 XX
 AC ABG61824;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Prostate cancer-associated protein #25.
 XX
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 XX

OS Mammalia.
 XX
 PN WO200230268-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US32045.
 XX
 PR 13-OCT-2000; 2000US-0687576.
 PR 08-DEC-2000; 2000US-0733288.
 PR 08-DEC-2000; 2000US-0733742.
 PR 24-JAN-2001; 2001US-263957P.
 PR 16-MAR-2001; 2001US-276791P.
 PR 16-MAR-2001; 2001US-276888P.
 PR 06-APR-2001; 2001US-281922P.
 PR 24-APR-2001; 2001US-286214P.
 PR 30-APR-2001; 2001US-0847046.
 PR 04-MAY-2001; 2001US-288589P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevez P;
 XX
 XX WPI: 2002-471335/50.
 DR N-PSDB; ABK92139.
 XX
 PT Detecting a prostate cancer-associated transcript in a cell in a
 PT patient, useful for diagnosing prostate cancer (PC) or screening
 PT modulators of PC, by determining if prostate cancer-associated genes
 PT are expressed in a prostate tissue -
 XX
 PS Claim 27; Page 321; 436pp; English.
 XX
 CC The present invention relates to methods of detecting a prostate
 CC cancer-associated transcript in a cell from a patient. The method
 CC comprises contacting a biological sample from the patient with
 CC prostate cancer-associated polynucleotides (designated PC genes) that
 CC selectively hybridise to a sequence that is at least 80% identical
 CC to them. The prostate cancer-associated polynucleotide sequences
 CC are differentially expressed in prostate tumour tissue or in
 CC prostate cancer and are derived from the tissues of various
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
 CC The methods of the invention are useful for diagnosing and treating
 CC prostate cancer in mammals. The prostate cancer-associated genes are
 CC useful for diagnosing or treating prostate cancer, as well as for
 CC identifying modulators of prostate cancer or agents that inhibit
 CC prostate cancer. The nucleic acid sequences are particularly useful
 CC in gene therapy as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
 CC
 SQ Sequence 1193 AA;
 XX
 QY 1 DLPMSPRGVIASNLF 16
 |||: |||: : :
 ID 1113 DLPMPRGFLPGHAF 1128
 DB
 RESULT 16
 AAU32407
 ID AAU32407 standard; Protein: 1194 AA.
 XX
 AC AAU32407;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #2898.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 XX

PT and also for testing and detection e.g. diagnosis -
 XX Claim 11: SEQ ID No 254; 546bp; English.
 XX
 CC The invention relates to novel isolated human prostate cancer antigen
 CC polynucleotides (I) and polypeptides (II) (I) and (II) are useful for
 CC preventing, treating or ameliorating a medical condition when
 CC administered, (I), (II) and the antibody to (II) are useful for treating,
 CC preventing and/or prognosing disorders related to the reproductive
 CC system including prostate cancers; urinary disorders e.g. chronic
 CC nephritis; and blood-related disorders e.g. thrombosis. (II) can be used
 CC for testing and detection e.g. as a chromosomal marker and in forensics.
 CC (I) and the anti-(II) antibody can be used in testing and detection in
 CC immunoassays. AAU22702-AAU22913 represent the human prostate cancer
 CC antigen amino acid sequences, and related amino acid sequences of the
 CC invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 84 AA:

Query Match 50.68; Score 39; DB 22; Length 84;
 Best Local Similarity 70.08; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPMSPRCVIA 11
 :|||
 DB 52 IPLSPAGVIA 61

RESULT 19

AAU94707
 ID AAU94707 standard; Protein: 84 AA.

AAU94707:

21-NOV-2001 (first entry)

Human reproductive system related antigen SEQ ID NO: 3365.

Human; reproductive system related antigen; reproductive system disorder;
 cancer; gene therapy.

Homo sapiens.

WO200155320-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US01339.

XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226581.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249309.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-46570/50.

N-PSDB; AAL00677.

Claim 11; SEQ ID NO 3365; 1297pp + Sequence Listing; English.
Is used in preventing, treating or ameliorating a medical condition

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a protein of the invention.

XX Sequence 84 AA;

Query Match

Best Local Similarity 50.6%; Score 39; DB 22; Length 84;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LPMSPRGVIA 11
: : : : :
Db 52 IPISPAGVIA 61

RESULT 20

AAM89886
ID AAM89886 standard; Protein; 97 AA.

XX AC AAM89886;
XX DT 07-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen SEQ ID NO:17479.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis.
OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 16-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231415.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232168.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.

```
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234957.
PR 25-SEP-2000; 2000US-0234958.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236357.
PR 29-SEP-2000; 2000US-0236358.
PR 29-SEP-2000; 2000US-0236359.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241185.
PR 20-OCT-2000; 2000US-0241186.
PR 20-OCT-2000; 2000US-0241187.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.

PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX DR WPI: 2001-483426/52.
XX DR N-PSDB: AAK62667.
XX
XX PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating cancers and
XX PT metastasis -
XX
XX PS Claim 11: SEQ ID NO 17479; 3071pp + Sequence Listing; English.
XX
XX CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
XX CC activity, and can be used in gene therapy and vaccine production. (I)
XX CC proteins and polynucleotides may be used in the prevention, diagnosis and
XX CC treatment of diseases associated with inappropriate (I) expression. For
XX CC example, they may be used to treat disorders associated with decreased
XX CC expression by rectifying mutations or deletions in a patient's genome
XX CC that affect the activity of (I) by expressing inactive proteins or to
XX CC supplement the patient's own production of (I). Additionally, (I)
XX CC polynucleotides may be used to produce the secreted (I), by inserting
XX CC the nucleic acids into a host cell and culturing the cell to express the
XX CC protein. (I) proteins and polynucleotides may be used to prevent,
XX CC diagnose and treat immune/hematopoietic-related diseases, especially
XX CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX CC to AAK67694 represent human immune/hematopoietic antigen genomic
XX CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
XX CC represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 97 AA:

Query Match 50.6%; Score 39; DB 22; Length 97;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 IPMSPRGVIA 11
DB 54 IPSPAGVIA 63

Search completed: January 2, 2003, 12:04:27
Job time : 37 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:03:04 ; Search time 34 Seconds

(without alignments)
13.846 Million cell updates/sec

Title: US-09-824-053-3

Sequence: 1 DLPMSRGVIASTLXF 16

Gapop 10.0, Gapext 0.5

262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/3A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/3B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/3A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/3B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/3A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/3B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	97.4	16	4	US-08-669-304-3
2	75	97.4	546	4	US-08-669-304-31
3	74	96.1	16	4	US-08-669-304-10
4	39	50.6	468	4	US-09-292-097-2
5	39	50.6	770	1	US-08-416-581B-1
6	39	50.6	770	1	US-08-416-581B-5
7	39	50.6	770	4	US-09-087-465-6
8	39	50.6	771	1	US-08-276-099A-14
9	39	50.6	771	1	US-08-781-890-14
10	38	49.4	393	4	US-09-387-418A-10
11	38	49.4	770	1	US-08-369-796-12
12	38	49.4	770	1	US-08-416-581B-9
13	38	49.4	770	2	US-08-852-091-12
14	38	49.4	770	3	US-08-820-754-12
15	38	49.4	770	3	US-08-956-652-12
16	38	49.4	770	3	US-08-956-869-12
17	38	49.4	770	3	US-09-012-710-8
18	38	49.4	770	3	US-08-948-547-12
19	38	49.4	770	4	US-09-364-970-3
20	38	49.4	770	4	US-09-364-970-5
21	38	49.4	770	4	US-09-556-273-8
22	38	49.4	770	4	US-08-956-653A-12
23	38	49.4	770	4	US-09-526-542-2
24	38	49.4	770	5	PCR-US95-17025-12
25	38	49.4	4544	1	US-08-469-486-52
26	38	49.4	4544	2	US-08-469-486-52
27	37	48.1	307	4	US-08-858-207A-481

28	37	48.1	607	2	US-08-472-534-5	Sequence 5, Appl1
29	37	48.1	616	4	US-09-134-001C-3646	Sequence 3646, Ap
30	37	48.1	861	1	US-08-484-105-18	Sequence 18, Appl
31	37	48.1	861	1	US-08-484-106-18	Sequence 18, Appl
32	37	48.1	2584	3	US-08-936-135-4	Sequence 4, Appl1
33	37	48.1	2588	3	US-08-936-135-2	Sequence 4, Appl1
34	36	46.8	366	4	US-09-134-001C-4799	Sequence 4799, Ap
35	36	46.8	564	4	US-09-211-704A-8	Sequence 8, Appl1
36	36	46.8	669	4	US-08-704-711A-3	Sequence 3, Appl1
37	36	46.8	669	4	US-09-521-220-3	Sequence 3, Appl1
38	36	46.8	669	4	US-09-391-104-29	Sequence 29, Appl
39	35	45.5	298	2	US-09-006-535-3	Sequence 3, Appl1
40	35	45.5	317	1	US-08-118-270-27	Sequence 27, Appl
41	35	45.5	317	5	PCR-US93-08528-27	Sequence 27, Appl
42	34	44.2	31	3	US-08-926-842B-58	Sequence 58, Appl
43	34	44.2	241	3	US-08-834-776A-2	Sequence 2, Appl1
44	34	44.2	244	1	US-08-289-699A-3	Sequence 3, Appl1
45	34	44.2	244	2	US-08-878-285-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-669-304-3
Sequence 3, Application US/08669304
Patent No. 6251626
GENERAL INFORMATION:
APPLICANT: Peter Strougard
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
TITLE OF INVENTION: METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hunton & Williams
STREET: 1900 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,304
FILING DATE: 12 July 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,910
FILING DATE: 7 June 1995
ATTORNEY/AGENT INFORMATION:
NAME: Stanislaus Akman
REGISTRATION NUMBER: 28,562
REFERENCE/DOCKET NUMBER: 320,000003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-1926
TELEFAX: (202) 778-2201
TELEX: No. 6251626e
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-669-304-3

Query Match 97.4%, Score 75; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 3; 1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

```
QY      1  DLPMSPRGVIASNLXF  16
        |||||
Db      1  DLPMSPRGVIASNLXF  16
```

RESULT 2
US-08-669-304--31

; Sequence 31, Application US/08669304
; patent No. 6251626

```
; GENERAL INFORMATION:
;
APPLICANT: Peter Stougaard
ADDRESS: 104 Oak Street
```

APPLICANT: Ole Cal Hansen

```

; TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
;
; TITLE OF INVENTION: METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME

```

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hulton & Williams
STREET: 1900 K Street, N.W.

COUNTRY: U.S.A.
ZIP: 20006-1109
COMPUTER READABLE FORM:

SOFTWARE: PatentIn Release #1.0, Version #1.25

MOLECULE TYPE: protein
; US-08-669-304-31

RESULT 3
US-08-669-304-10

; Sequence 10, Application US/08669304

; Patent No. 6251626
; GENERAL INFORMATION:

APPLICANT: Peter Stougaard

APPLICANT: Ole Cai Hansen
TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A

TITLE OF INVENTION: MET

NUMBER OF SEQUENCES: 34

NUMBER OF REFERENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hunton & Williams

STREET: 1900 K S
CITY: Washington

```

; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006-1109
; COMPUTER READABLE
;

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; PATENTROW NUMBER: 76/08/650 204

```

MOLECULE TYPE: peptide
US-08-669-304-10

Best Local Similarity	93.8%	Pred. NO. 4.6e-07;			
Matches	15; Conservative	0; Mismatches	1; Indels	0; Gaps	0;

```

QY      1 DLPMSPRGVIASNLXF 16
        |||||

```

OTHER INFORMATION: 103348CD1
US-09-292-097-2

Best Local Similarity	37.58;	Pred. No. 39;
Matches	6; Conservative	4; Indels
		0; Gaps

QY 1 DLPMSPRGVIASNLXF 16
:|:|:|:|:|:|

```

      176 NLPLSPQGTVRTAVEEF 191
Db

```

RESULT 5

US-08-416-581B-1
; Sequence 1, Application US/08416581B
; Patent No. 5719042
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadimitsu
; APPLICANT: Akira, Shizuo
; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,581B
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-65825/1994
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: O-37891
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-581B-1

Query Match 50.6%; Score 39; DB 1; Length 770;
Best Local Similarity 77.8%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 DLPMSPRGV 9
|||||:
723 DLPMSPRAL 731

RESULT 6
US-08-416-581B-5
; Sequence 5, Application US/08416581B
; Patent No. 5719042
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadimitsu
; APPLICANT: Akira, Shizuo
; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,581B
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-65825/1994
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: O-37891
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-581B-5

Query Match 50.6%; Score 39; DB 1; Length 770;
Best Local Similarity 77.8%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 DLPMSPRGV 9
|||||:
723 DLPMSPRAL 731

RESULT 7
US-09-87-465-6
; Sequence 6, Application US/09087465A
; Patent No. 6160092
; GENERAL INFORMATION:
; APPLICANT: Vinkemeler, Uwe
; APPLICANT: Chen, Xiaomlin
; APPLICANT: Darnell Jr., James E
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: 600-1-229
; CURRENT APPLICATION NUMBER: US/09/087,465A
; CURRENT FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-87-465-6

Query Match 50.6%; Score 39; DB 4; Length 770;
Best Local Similarity 77.8%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 DLPMSPRGV 9
|||||:
723 DLPMSPRAL 731

RESULT 8
US-08-276-099A-14
; Sequence 14, Application US/08276099A
; Patent No. 5591825
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L
; APPLICANT: Hou, Jinhao
; TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
; TITLE OF INVENTION: BINDING ASSAYS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,099A
FILING DATE: 15-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59451-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-276-099A-14

Query Match 50.6%; Score 39; DB 1; Length 771;
Best Local Similarity 77.8%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLPMSRGV 9
Db 723 DLPMSPRAL 731

RESULT 9
US-08-781-890-14
Sequence 14, Application US/08781890
Patent No. 5710266
GENERAL INFORMATION:
APPLICANT: McKnight, Steven L
APPLICANT: Hou, Jinzhao
TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
NUMBER OF SEQUENCES: 17
TITLE OF INVENTION: BINDING ASSAYS
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,890
FILING DATE: 05-JAN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,099
FILING DATE: 15-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59451-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-781-890-14

Query Match 50.6%; Score 39; DB 1; Length 771;
Best Local Similarity 77.8%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLPMSRGV 9
Db 723 DLPMSPRAL 731

RESULT 10
US-09-387-418A-10
Sequence 10, Application US/09387418A
Patent No. 6391572
GENERAL INFORMATION:
APPLICANT: Zhang, Xiaokui
APPLICANT: Wreszczynska, Melissa H
APPLICANT: Horvath, Curt M
APPLICANT: Darnell Jr., James E
TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVAT
FILE REFERENCE: 600-1-253
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10
LENGTH: 393
TYPE: PRT
ORGANISM: Mus musculus
US-09-387-418A-10

Query Match 49.4%; Score 38; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPR 7
Db 346 DLPMSPR 352

RESULT 11
US-08-369-796-12
Sequence 12, Application US/08369796
Patent No. 5716622
GENERAL INFORMATION:
APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong
TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

```
COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-369-796-12

Query Match          49.4%; Score 38; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPR 7
DB 723 DLPMSPR 729

RESULT 12
; Sequence 9, Application US/08416581B
; Patent No. 5719042
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadamiatsu
; APPLICANT: Akira, Shizuo
; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,581B
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-65825/1994
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: 0-37891
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
```

```
TOPLOGY: linear
; MOLECULE TYPE: protein
; US-08-416-581B-9

Query Match          49.4%; Score 38; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPR 7
DB 723 DLPMSPR 729

RESULT 13
; Sequence 12, Application US/08852091
; Patent No. 5883228
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-091-12

Query Match          49.4%; Score 38; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPR 7
DB 723 DLPMSPR 729

RESULT 14
; Sequence 12, Application US/08820754
; Patent No. 5976835
```

GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820,754
FILING DATE: 19-MAR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-956-652-12

Query Match 49.4%; Score 38; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e+02; 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSR 7
DB 723 DLPMSR 729

RESULT 15
US-08-956-652-12
Sequence 12, Application US/08956652
Patent No. 6013475
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong

APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,652
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-956-652-12

Query Match 49.4%; Score 38; DB 3; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e+02; 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSR 7
DB 723 DLPMSR 729

RESULT 16
US-08-956-869-12
Sequence 12, Application US/08956869
Patent No. 6030808
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,869
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/212,185
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-956-869-12

Query Match
Best Local Similarity 49.4%; Score 38; DB 3; Length 770;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLPMSR 7
723 DLPMSR 729

RESULT 17
US-09-012-710-8
Sequence 8, Application US/09012710
Patent No. 6087478
GENERAL INFORMATION:
APPLICANT: Vinkemeter, Uwe
APPLICANT: Moarefi, Ismail
APPLICANT: Darnell, Jr., James E.
APPLICANT: Kurlivan, John
TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,710
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-194
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-012-710-8

Query Match
Best Local Similarity 49.4%; Score 38; DB 3; Length 770;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSR 7
Db 723 DLPMSR 729

RESULT 18
US-08-948-547-12
Sequence 12, Application US/08948547
Patent No. 6124118
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,547
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569

FILING DATE: 19-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/126,588
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-073 CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEFAX: 133521
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-948-547-12

Query Match 49.4%; Score 38; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 DLPMSPR 7
 DB 723 DLPMSPR 729

RESULT 19
 US-09-364-970-3
 Sequence 3, Application US/09364970
 Patent No. 6235873
 GENERAL INFORMATION:
 APPLICANT: Bromberg, Jacqueline
 TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
 TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
 TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
 FILE REFERENCE: 600-1-252
 CURRENT APPLICATION NUMBER: US/09/364,970
 CURRENT FILING DATE: 1999-07-31
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 3
 LENGTH: 770
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-364-970-3

Query Match 49.4%; Score 38; DB 4; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 DLPMSPR 7
 DB 723 DLPMSPR 729

RESULT 20
 US-09-364-970-5
 Sequence 5, Application US/09364970
 Patent No. 6235873
 GENERAL INFORMATION:
 APPLICANT: Bromberg, Jacqueline
 TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
 TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
 TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
 FILE REFERENCE: 600-1-252
 CURRENT APPLICATION NUMBER: US/09/364,970
 CURRENT FILING DATE: 1999-07-31
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 5

LENGTH: 770
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-364-970-5
 Query Match 49.4%; Score 38; DB 4; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 DLPMSPR 7
 DB 723 DLPMSPR 729
 Search completed: January 2, 2003, 12:05:31
 Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:03:04 ; Search time 16 Seconds

(without alignments)
96.134 Million cell updates/sec

Title: US-09-824-053-3

Perfect score: 77
Sequence: 1 DLPMSPRGVASNLXF 16

Listing table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Detected: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	55.8	184	2	E98229
2	43	55.8	184	2	AI3056
3	42	54.5	416	2	B65171
4	41	53.2	416	2	E91207
5	41	53.2	416	2	G86053
6	40	51.9	295	2	A60131
7	40	51.9	407	2	G84309
8	39	50.6	113	2	I51025
9	39	50.6	206	2	A64339
10	39	50.6	770	2	A54444
11	38	49.4	272	1	C64009
12	38	49.4	328	2	E95421
13	38	49.4	334	2	E86597
14	38	49.4	334	2	G72024
15	38	49.4	363	2	G72854
16	38	49.4	479	2	S72582
17	38	49.4	670	2	A75542
18	38	49.4	684	2	G90473
19	38	49.4	770	2	I49508
20	38	49.4	1305	2	T31096
21	38	49.4	4543	1	A53102
22	38	49.4	4544	1	S02392
23	38	49.4	4545	1	S25111
24	37	48.1	196	2	T50018
25	37	48.1	245	2	AB1807
26	37	48.1	362	2	T41785
27	37	48.1	368	2	G75425
28	37	48.1	485	1	F71375
29	37	48.1	500	2	C84791

30	37	48.1	605	2	I39837	dnak-type molecule
31	37	48.1	607	2	B86744	dnak protein (limpo
32	37	48.1	607	2	S39342	dnak-type molecule
33	37	48.1	607	2	B95060	dnak protein (limpo
34	37	48.1	607	2	G97928	hypothetical prote
35	37	48.1	611	2	S09500	dnak-type molecule
36	37	48.1	614	2	B83818	class I heat-shock
37	37	48.1	631	2	S71508	dnak-type molecule
38	37	48.1	637	2	B83052	dnak protein PA476
39	37	48.1	749	2	A99438	hypothetical prote
40	37	48.1	860	2	A43920	nucleolar 100K pol
41	37	48.1	861	2	G02329	replication contro
42	37	48.1	885	2	JH0796	nucleolar 100K pol
43	37	48.1	1199	1	S76349	transcription-repa
44	37	48.1	1477	2	T18534	protein-tyrosine k
45	37	48.1	2218	2	B84683	hypothetical prote

ALIGNMENTS

RESULT 1

E98229

sarcosine oxidase gamma chain PA5419 [imported] - Agrobacterium tumefaciens (strain C

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: E98229

R:Goodner, B.; Hinkle, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Quirolo, B.; Goldm

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E98229

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-184 <KUR>

A:Cross-references: GB:AE007870; PIDN:AKR9359.1; PID:gl5159207; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L_1572

A:Map position: linear chromosome

C:Superfamily: Corynebacterium sp. sarcosine oxidase gamma chain

Query Match	55.8%	Score 43;	DB 2;	Length 184;
Best Local Similarity	53.3%	Pred. No. 3.2;		
Matches	8;	Conservative	2;	Mismatches
			5;	Indels
				Gaps
				0;

QY 2 LPMSPRGVIANLXF 16

DB 54 LPTSPKGSVTSGLRF 68

RESULT 2

AI3056

sarcosine oxidase gamma subunit [imported] - Agrobacterium tumefaciens (strain C58, D

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AI3056

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl

i; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Teo, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AI3056

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-184 <KUR>

A:Cross-references: GB:AE006889; PIDN:AAL4871.1; PID:gl7742519; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: soxg

RESULT 5

Riñán, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja Jung, K.H.; Alam, M.; Freitas, T.

C:Accession: E95421
 R: Barnett, M.T.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: E95421
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-328 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK65935.1; PID:g14524450; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Flatau, T.M.; Long, S.R.; Fuhrer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 P.; Chailin, P.; Cowle, A.; Davits, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 Hubault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Genetics: annotation
 A:Gene: gpda
 A:Superfamily: plasmid

Query Match 49.4%; Score 38; DB 2; Length 328;
 Best Local Similarity 64.3%; Pred. No. 49;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 DLPSPRGVIASNL 14
 DB 4 DLMSSRRVIALSGL 17

RESULT 13
 E86597
 glycerol-3-P dehydrogenase [imported] - *Chlamydomonas reinhardtii* (strain J138)
 C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: E86597
 R:Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of *Chlamydomonas reinhardtii* J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: E86597
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <STO>
 A:Cross-references: GB:BA000008; NID:g8979229; PIDN:BA99063.1; GSPDB:GN00142
 A:Experimental source: strain J138
 A:Genetics:
 A:Gene: gpda
 A:Superfamily: glycerol-3-phosphate dehydrogenase

Query Match 49.4%; Score 38; DB 2; Length 334;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 PMSPRGVIASNL 16
 DB 48 PLAPNVVISPLNSF 61

RESULT 14
 G72024
 glycerol-3-phosphate dehydrogenase, NAD-dependent CP1014 [imported] - *Chlamydomonas reinhardtii*
 C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: G72024; D81513
 R:Kalman, S.; Mitchell, W.; Marathe, R.; Lamm, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of *Chlamydomonas reinhardtii* and *C. trichomonas*.
 A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: G72024
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <CAR>
 A:Cross-references: GB:AE001667; GB:AE001363; NID:g4377171; PIDN:AA018993.1; PID:g437
 A:Experimental source: strain CWM1029
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of *Chlamydomonas reinhardtii* MOPn and *Chlamydomonas reinhardtii* AR39
 A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: D81513
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <REA>
 A:Cross-references: GB:AE002258; GB:AE002161; NID:g7189919; PIDN:AAF38792.1; PID:g718
 A:Experimental source: strain AR39, HL cells
 A:Genetics:
 A:Gene: gpda; CP1014
 A:Superfamily: glycerol-3-phosphate dehydrogenase

Query Match 49.4%; Score 38; DB 2; Length 334;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 PMSPRGVIASNL 16
 DB 48 PLAPNVVISPLNSF 61

RESULT 15
 G72854
 hypothetical protein - Autographa californica nuclear polyhedrosis virus
 C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
 A:Note: dsDNA virus
 C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 21-Jul-2000
 C:Accession: G72854; A45684
 R:Atres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Posselt, R.D.
 Virology 202, 586-605, 1994
 A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
 A:Reference number: A72850; MUID:94303173; PMID:8030224
 A:Accession: G72854
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-363 <ATV>
 A:Cross-references: GB:U22658; NID:g510708; PIDN:AAA66669.1; PID:g559108
 R:Carstens, E.B.; Lu, A.L.; Chan, H.L.
 J. Virol. 67, 2513-2520, 1993
 A:Title: Sequence, transcriptional mapping, and overexpression of p47, a baculovirus
 A:Reference number: A45684; MUID:93233214; PMID:8474157
 A:Accession: A45684
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-363 <CAR>
 A:Cross-references: GB:U07878; NID:g293990; PIDN:AAA16858.1; PID:g293991
 C:Genetics:
 A:Gene: Ac-p43

Query Match 49.4%; Score 38; DB 2; Length 363;
 Best Local Similarity 40.0%; Pred. No. 54;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 LPMSPRGVIASNL 16
 DB 183 MEVKKRGGVKSNTAF 197

RESULT 16
 S72582
 hypothetical protein B1937_C3_231 - *Mycobacterium leprae*
 C:Species: *Mycobacterium leprae*
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 22-Oct-1999

C:Accession: S72582
 R:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S72580
 A:Accession: S72582
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-479 <SMI>
 A:Cross-references: EMBL:U00016; NID:g466961; PIDN:AAAI1150.1; PID:g466964
 C:Genetics:
 A:Start codon: GTG

Query Match 49.4%; Score 38; DB 2; Length 479;
 Best Local Similarity 70.0%; Pred. No. 75;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 2 LPMSPRGVIA 11
 ||| ||| |||
 225 LPLGPRGVNA 234

RESULT 17
 A75542
 conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: A75542
 R:White, O.; Eelsen, J.A.; Heldelberg, J.F.; Hickay, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75540; MUID:20036896; PMID:10567266
 A:Accession: A75542
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-670 <WHI>
 A:Cross-references: GB:AE00186; GB:AE000513; NID:g6457921; PIDN:AAF09837.1; PID:g645792
 C:Genetics:
 A:Experimental source: strain R1
 A:Gene: DR0250
 A:Map position: 1

Query Match 49.4%; Score 38; DB 2; Length 670;
 Best Local Similarity 58.3%; Pred. No. 1.1e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 5 SPRGVASNLKF 16
 ||| ||| ||| |||
 555 TPAGVASSLIF 566

RESULT 18
 G90473
 hypothetical protein cutA-6 [imported] - *Sulfolobus solfataricus*
 C:Species: *Sulfolobus solfataricus*
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-Aug-2001
 C:Accession: G90473
 R:Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
 arett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: *Sulfolobus solfataricus* complete genome.
 A:Reference number: A99139
 A:Accession: G90473
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-684 <KUR>
 A:Cross-references: GB:AE006641; NID:g13816315; PIDN:AAK43046.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: cutA-6
 C:Superfamily: carbon-monoxide dehydrogenase large chain

Query Match 49.4%; Score 38; DB 2; Length 684;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 3 PMSPRGVIA 11
 ||| ||| |||
 176 PMEPKGIIA 184

RESULT 19
 I49508
 ISGF3 p91-related transcription factor - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
 C:Accession: I49508; MUID:149009
 R:Kira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.;
 Cell 77, 63-71, 1994
 A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related
 A:Reference number: A54444; MUID:94208062; PMID:7512451
 A:Accession: I49508
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-770 <RES>
 A:Cross-references: GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g476716
 R:Raz, R.; Durbin, J.E.; Levy, D.E.
 J. Biol. Chem. 269, 24391-24395, 1994
 A:Title: Acute phase response factor and additional members of the interferon-stimula
 A:Reference number: I49009; MUID:95014185; PMID:7523373
 A:Accession: I49009
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-393; MUID:395-700; 702-770 <RE2>
 A:Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473889
 C:Genetics:
 A:Gene: APRF
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 49.4%; Score 38; DB 2; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 DLPMSPR 7
 ||| ||| |||
 723 DLPMSPR 729

RESULT 20
 T31096
 cyclin G-associated kinase GAK - rat
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T31096
 R:Kanaoka, Y.; Kimura, S.H.; Okazaki, I.; Ikeda, M.; Nojima, H.
 FEBS Lett. 402, 73-80, 1997
 A:Title: GAK: a cyclin G-associated kinase contains a tensin/auxilin-like domain.
 A:Reference number: Z20979; MUID:97165969; PMID:9013862
 A:Accession: T31096
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1305 <KAN>
 A:Cross-references: EMBL:D38560; NID:g1902912; PIDN:BAI18911.1; PID:g1902913
 A:Note: GAK and cyclin G associate together in vivo
 C:Genetics:
 A:Gene: GAK

Query Match 49.4%; Score 38; DB 2; Length 1305;
 Best Local Similarity 61.5%; Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 1 DLPMSPRGVASN 13
 ||| ||| ||| |||
 1015 DLPMPKVNASS 1027

Thu Jan 2 12:35:48 2003

us-09-824-053-3.rpt

Page 6

Search completed: January 2, 2003, 12:04:49
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:03:00 ; Search time 32 Seconds

(without alignments)
20.738 Million cell updates/sec

Title: US-09-824-053-3
Sequence: 1 DLPMSRCVIA5NLXF 16

Align table: BLOSUM62
Gapop 10.0, Gapext 0.5

Database: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	54.5	416	1 YIDR_ECOLI	P31455 escherichia
2	40	51.9	295	1 HX71_XENLA	Q04281 xenopus lae
3	39	50.6	206	1 Y312_METJA	Q57760 methanococ
4	39	50.6	770	1 STR3_HUMAN	P40763 homo sapien
5	38	49.4	272	1 RN26_HAEIN	P44012 haemophilus
6	38	49.4	334	1 GPDA_CHLPN	Q92751 chlamydia p
7	38	49.4	363	1 VP43_NPVAC	P34050 autographa
8	38	49.4	590	1 DNK3_STRMO	Q06942 streptococ
9	38	49.4	770	1 STR3_MOUSE	P42227 mus musculu
10	38	49.4	770	1 STR3_RAT	P52631 rattus norv
11	38	49.4	1305	1 GAK_RAT	P97874 rattus norv
12	38	49.4	4543	1 LRP1_CHICK	P98157 gallus gall
13	38	49.4	4544	1 LRP1_HUMAN	Q07954 homo sapien
14	37	48.1	196	1 FLC_ARATH	Q98797 arabidopsis
15	37	48.1	499	1 GAK_MOUSE	Q09474 mus musculu
16	37	48.1	500	1 GAPS_PSEAG	Q93173 pseudomonas
17	37	48.1	604	1 DNK3_BACME	P05646 bacillus me
18	37	48.1	606	1 DNK3_BACTR	Q9KWS7 bacillus th
19	37	48.1	607	1 DNK3_LACLA	P42368 lactococcus
20	37	48.1	607	1 DNK3_STRPN	P95829 streptococ
21	37	48.1	609	1 DNK3_STRPY	P95831 streptococ
22	37	48.1	609	1 DNK3_STRAG	P95693 streptococ
23	37	48.1	610	1 DNK3_BACSH	Q09268 bacillus sp
24	37	48.1	610	1 DNK3_BACSU	P17620 bacillus su
25	37	48.1	613	1 DNK3_BACBD	Q9K472 bacillus ha
26	37	48.1	613	1 DNK3_LACSK	Q87777 pseudocall
27	37	48.1	637	1 DNK3_PSEAE	Q9H443 pseudomonas
28	37	48.1	644	1 DNK3_LEGPN	Q32482 legionella
29	37	48.1	656	1 DNK3_COXBU	Q87712 coxiella bu
30	37	48.1	861	1 ORC1_HUMAN	Q13415 homo sapien
31	37	48.1	885	1 PWC2_HUMAN	Q01180 homo sapien
32	37	48.1	887	1 PWC2_MOUSE	P56660 mus musculu
33	37	48.1	1199	1 MFD_SYNY3	Q55750 synchocyst

34	37	48.1	1377	1 NEOL_RAT	P97603 rattus norv
35	37	48.1	1461	1 NEOL_HUMAN	Q92859 homo sapien
36	37	48.1	1477	1 HTK7_HYDAT	Q25197 hydra atten
37	37	48.1	1493	1 NEOL_MOUSE	P97798 mus musculu
38	37	48.1	1493	1 Y698_CHLMU	Q9P188 chlamydia m
39	36.5	47.4	293	1 YC76_MYCTU	Q11043 mycobacteri
40	36	46.8	158	1 YC76_MYCTU	Q11043 mycobacteri
41	36	46.8	269	1 NUDC_VIRCH	Q9Kv27 vibrio chol
42	36	46.8	349	1 PECM_ERWCH	Q07251 aicalligenes
43	36	46.8	388	1 LDH_AICEU	P45460 versinia en
44	36	46.8	409	1 AMPC_YEREN	P17778 versinia pe
45	36	46.8	421	1 Y118_TREPA	Q83155 treponema p

ALIGNMENTS

RESULT 1					
YIDR_ECOLI	STANDARD:	PRT:	416 AA.		
AC P31455; P76734;					
DT 01-JUL-1993 (Rel. 26, Created)					
DT 15-JUL-1998 (Rel. 36, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE Hypothetical protein yidr.					
GN YIDR OR B3689.					
OS Escherichia coli.					
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;					
OX NCBI_TaxID=562;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=K12 / MG1655;					
RX MEDLINE=93315143; PubMed=7686882;					
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;					
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli					
RL Genomics 16:551-561(1993).					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
DR EMBL: L10328; AAA62041.1; -					
DR EMBL: AEO00446; AAC76712.1; -					
DR Ecogene; EG11713; yidr.					
KW Hypothetical protein; Complete proteome.					
SQ SEQUENCE 416 AA; 46318 MW; 47E36360CC89A8B7 CRC64;					
Query Match	54.5%	Score 42;	DB 1;	Length 416;	
Best Local Similarity	53.3%	Pred. No. 4.9;			
Matches 8;	Conservative 1;	Mismatches 6;	Indels 0;	Gaps 0;	
QY 2 LPMSPRCVIA5NLXF 16					
Db 272 LPAPRGVGRRLTF 286					
RESULT 2					
HX71_XENLA	STANDARD:	PRT:	295 AA.		
ID HX71_XENLA					
AC C04281;					
DT 01-JUN-1994 (Rel. 29, Created)					
DT 01-JUN-1994 (Rel. 29, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE Homeobox protein XHOX-7.1 (Fragment).					
GN XHOX-7.1.					
OS Xenopus laevis (African clawed frog).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: L29277; AAS58374.1; -
 DR HSSP: P42227; IBC1.
 DR TRANSFAC: T01493; -
 DR Genew: HGNC:11364; STAT3.
 DR MIM: 102582; -
 DR InterPro: IPR00980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR Pfam: PF02864; STAT_prot; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.
 FT DOMAIN 580 670
 FT MOD_RES 705 705 PHOSPHORYLATION (BY JAK) (BY SIMILARITY).
 FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
 SO SEQUENCE 770 AA; 88052 MW; A3DCE815B3B5360 CRC64;

Query Match 50.6%; Score 39; DB 1; Length 770;
 Best Local Similarity 77.8%; Pred. No. 34;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DLPMSRGV 9
 Db 723 DLPMSRPL 731

RESULT 5
 RN26_HAFTN STANDARD; PRT; 272 AA.
 AC P44012;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable ribonuclease HI0526 precursor.
 GN HI0526.
 OS Haemophilus influenzae.
 BA Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 GN Haemophilus.
 NCBI_TaxID=727;

SEQUENCE FROM N.A.
 RC STRAIN-Rd / RW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Flaischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA McEwen A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 CC -1- SIMILARITY: BELONGS TO THE RNASE T2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: U32735; AAC22192.1; -
 DR TIGR: HI0526; -
 DR InterPro: IPR001568; RNase_T2.
 DR Pfam: PF00445; ribonuclease_T2; 1.
 DR PROSITE: PS00530; RNase_T2_1; 1.
 DR PROSITE: PS00531; RNase_T2_2; 1.
 KW Hypothetical protein; Hydrolase; Nuclease; Endonuclease; Signal;
 FW Complete proteome.
 FT SIGNAL 1 23
 FT CHAIN 24 272
 FT ACT_SITE 148 148
 FT ACT_SITE 195 195
 FT ACT_SITE 199 199
 FT ACT_SITE 199 199
 SO SEQUENCE 272 AA; 31482 MW; 9D50F9C2CEB392 CRC64;

Query Match 49.4%; Score 38; DB 1; Length 272;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 DLPMSRGVIASNL 14
 Db 170 DLPALPKGLAQL 183

RESULT 6
 GPDA_CHLPN STANDARD; PRT; 334 AA.
 AC Q92751;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94) (NAD(P)H-
 DE dependent glycerol-3-phosphate dehydrogenase).
 GN GPDA OR GPDA OR CPN0855 OR CP1014.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 GN NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CW1029.
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Liner K., Weidman J., Khoult H., Craven B., Bowman C., Dodson R.,
 RA Gwin W., Nelson W., Deboy R., Koloney J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
 pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kubara S., Nakazawa T.;
 "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CW1029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P)(+) =
 CC glycerone phosphate + NAD(P)H.
 CC -1- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
 CC formation.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE001667; AAD18993.1; -
 CC EMBL: AE002258; AAF8792.1; -
 CC EMBL: AP002548; BAA99063.1; -
 CC TIGR: CP1014; -
 CC InterPro: IPR001652; NAD_Gly3P-dh.
 CC Pfam: PF01210; NAD_Gly3P-dh; 1.
 CC PRINTS: PR00077; GPDHRCGNASE.
 CC ProDom: PD001649; NAD_Gly3P-dh; 1.
 CC PROSITE: PS00957; NAD_G3PDH; 1.
 CC Phospholipid biosynthesis; Oxidoreductase; NAD: Complete proteome.
 CC SEQUENCE 334 AA; 36161 MW; 4339ED6FFD77CEB6 CRC64;
 CC
 CC Query Match 49.4%; Score 38; DB 1; Length 334;
 CC Best Local Similarity 50.0%; Pred. No. 20;
 CC Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 3 PMSPRGVASNLX 16
 CC Db 48 PLAVNVVISPLSLF 61
 CC
 CC RESULT 7
 CC ID VP43_NPVAC STANDARD; PRT; 363 AA.
 CC AC P34050;
 CC DT 01-FEB-1994 (Rel. 28, Created)
 CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
 CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
 CC DE 43 kDa protein.
 CC GN P43.
 CC OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 CC OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC NCBI_TaxID=46015;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=HR3;
 CC RX MEDLINE=9323214; PubMed=8474157;
 CC RA Carstens E.B., Lu A.L., Chan H.L.B.;
 CC "Sequence, transcriptional mapping, and overexpression of p47, a
 CC baculovirus gene regulating late gene expression.";
 CC J. Virol. 67:2513-2520(1993).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=C6;
 CC RX MEDLINE=9430173; PubMed=8030224;
 CC RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 CC "The complete DNA sequence of Autographa californica nuclear
 CC polyhedrosis virus.";
 CC RT Virology 202:586-605(1994).
 CC RL
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L07878; AAA16858.1; -
 CC EMBL: L22858; AAA66669.1; -

DR PIR: A45684; A45684.
 SO SEQUENCE 363 AA; 43490 MW; DCBD20B50B078736 CRC64;
 CC
 CC Query Match 49.4%; Score 38; DB 1; Length 363;
 CC Best Local Similarity 40.0%; Pred. No. 22;
 CC Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC QY 2 LPMSPRGVASNLX 16
 CC Db 183 MPVKKKGVKSNYAF 197
 CC
 CC RESULT 8
 CC ID DNK STRMU STANDARD; PRT; 590 AA.
 CC AC 006942;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Chaperone protein dnak (Heat shock protein 70) (Heat shock 70 kDa
 CC protein) (Hsp70) (Fragment).
 CC GN DNK.
 CC OS Streptococcus mutans.
 CC OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC OC Streptococcus.
 CC NCBI_TaxID=1309;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=GS-5;
 CC RX MEDLINE=97426041; PubMed=9282745;
 CC RA Jayaraman G.C., Penders J.E., Burne R.A.;
 CC "Transcriptional analysis of the Streptococcus mutans hrca, gripe and
 CC dnak genes and regulation of expression in response to heat shock and
 CC environmental acidification.";
 CC RT Mol. Microbiol. 25:329-341(1997).
 CC RL
 CC CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC CC -1- INDUCTION: BY HEAT SHOCK.
 CC CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U78296; AAC45612.1; -
 CC DR HSSP: P04475; 2BPR.
 CC DR InterPro: IPR001023; Hsp70.
 CC DR Pfam: PF00012; HSP70; 1.
 CC DR ProDom: PD000089; Hsp70; 1.
 CC DR PROSITE: PS00297; HSP70_1; 1.
 CC DR PROSITE: PS00329; HSP70_2; 1.
 CC DR PROSITE: PS01036; HSP70_3; 1.
 CC KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
 CC FT MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CC FT NON_TER 590 590
 CC SO SEQUENCE 590 AA; 63258 MW; 90787ADF08F702B9 CRC64;
 CC
 CC Query Match 49.4%; Score 38; DB 1; Length 590;
 CC Best Local Similarity 43.8%; Pred. No. 38;
 CC Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 CC
 CC QY 1 DLPMSPRGVASNLX 16
 CC Db 432 DIPAPRGVQLIEYF 447
 CC
 CC RESULT 9
 CC ID STA3_MOUSE STANDARD; PRT; 770 AA.
 CC AC P42227;

DF 01-NOV-1995 (Rel. 32, Created)
 DF 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Signal transducer and activator of transcription 3 (Acute-phase response factor).
 DE STAT3 OR APF.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 154-158; 181-185 AND 632-640.
 RC STRAIN-BALB/C; TISSUE=Liver;
 RX MEDLINE-94208062; PubMed-7512451;
 AK Arita S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T., Yoshida K., Sudo T., Naruto M., Kishimoto T.;
 RT Molecular cloning of APF, a novel IFN-gamma-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway.";
 RL Cell 77:63-71(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE-94188718; PubMed-8140422;
 RA Zhong Z., Wen Z., Darnell J.E. Jr.;
 RT "Stat3, a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and Interleukin-6.";
 RL Science 264:95-98(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-95014185; PubMed-7523373;
 RA Raz R., Durbin J.E., Levy D.E.;
 RT "Acute phase response factor and additional members of the interferon-gamma-stimulated gene factor 3 family integrate diverse signals from cytokines, interferons, and growth factors.";
 RL J. Biol. Chem. 269:24391-24395(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (STAT3B).
 RC STRAIN-BALB/C; and C57BL/6; TISSUE=Liver;
 RX MEDLINE-96016116; PubMed-7568080;
 RA Schaefer T.S., Sanders L.K., Nathans D.;
 RT "Cooperative transcriptional activity of Jun and Stat3 beta, a short form of Stat3.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
 RN [5]
 RP PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.
 RC MEDLINE-95354205; PubMed-7543024;
 RT Wen Z., Zhong Z., Darnell J.E. Jr.;
 RL "Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation.";
 RL Cell 82:241-250(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
 RX MEDLINE-98334373; PubMed-9671298;
 RA Becker S., Groner B., Mueller C.W.;
 RT "Three-dimensional structure of the Stat3beta homodimer bound to DNA.";
 RL Nature 394:145-151(1998).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6 (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS ACUTE-PHASE PROTEIN GENES. STAT3B INTERACTS WITH THE N-TERMINAL PART OF C-JUN TO ACTIVATE SUCH PROMOTERS IN A COOPERATIVE WAY.
 CC -1- PATHWAY: INVOLVED IN THE GP130-MEDIATED SIGNALING PATHWAY.
 CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY MEMBER (AT LEAST STAT1) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear; translocated into the nucleus in response to phosphorylation.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: STAT3A (SHOWN HERE) AND STAT3B; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: STAT3A IS SEEN IN THE LIVER, SPLEEN, AND KIDNEY. STAT3B IS ALSO DETECTED IN THE LIVER, ALTHOUGH IN A MUCH LESS ABUNDANT MANNER.

CC -1- P.TM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF, LIF, CSE-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3 HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: L29278; AAA37254.1; -
 CC EMBL: 006922; AAA19452.1; -
 CC EMBL: 008378; AAA56668.1; -
 CC EMBL: 030709; AAC52612.1; -
 CC PDB: 1B61; 13-JAN-99.
 CC TRANSFAC: T01574; -
 CC MGD: MGI:103038; Stat3.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001217; STAT.
 CC Pfam: PF00017; SH2; 1.
 CC Pfam: PF01017; STAT; 1.
 CC Pfam: PF02864; STAT_bind; 1.
 CC Pfam: PF02865; STAT_prot; 1.
 CC SMART: SM00252; SH2; 1.
 CC PROSITE: PS50001; SH2; 1.
 CC Transcription regulation; DNA-binding; Nuclear protein; SH2 domain; Phosphorylation; Alternative splicing; 3D-structure.
 CC DOMAIN 580 670
 CC MOD_RES 705 705
 CC MOD_RES 727 727
 CC VARSPIC 716 770
 CC VARIANT 701 701
 CC MUTAGEN 727 727
 CC CONFLICT 16 16
 CC CONFLICT 25 25
 CC CONFLICT 394 394
 CC SEQUENCE 770 AA; 88053 MW; 6C00626711C8012D CRC64;
 Query Match 49.4%; Score 38; DB 1; Length 770;
 Best local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLPMSR 7
 DB 723 DLPMSR 729
 RESULT 10
 STR3_RAT
 ID STR3_RAT STANDARD: PRT; 770 AA.
 AC P52631;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Signal transducer and activator of transcription 3.
 GN STAT3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE-96102059; PubMed-8530402;
 RA Ripberger J.A., Filtz S., Richter K., Hocke G.M., Lottspeich F.,

RA Fey G.H.;
RT "transcription factors Stat3 and Stat5b are present in rat liver
nuclei late in an acute phase response and bind interleukin-6
response elements.";
RL J. Biol. Chem. 270:29998-30006(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
(IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
CC ACUTE-PHASE PROTEIN GENES.
CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
CC MEMBER (AT LEAST STAT1) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear; translocated into the nucleus in
CC response to phosphorylation (By similarity).
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAG3
CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X91810; CAA62920.1; -
DR HSSP: P42227; 1BGL
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR Pfam: PF02864; STAT bind; 1.
DR Pfam: PF02865; STAT_prot; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
DR Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 580 670 SH2.
FT MOD_RES 705 705 PHOSPHORYLATION (BY JAK) (BY SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 770 AA: 88039 MW: D74A0C76954754ED CRC64;

Query Match 49.4%; Score 38; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLPMSPR 7
DB 723 DLPMSPR 729

RESULT 11

RA GAK RAT
RT GAK RAT
RL GAK RAT
CC GAK RAT
CC 01-NOV-1997 (Rel. 35, Created)
CC 16-OCT-2001 (Rel. 40, Last annotation update)
CC Cyclin G-associated kinase (EC 2.7.1.-).
CC GAK.
CC OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=Kidney;
CC MEDLINE=97165969; PubMed=9013862;
CC Kanaka Y., Kimura S.H., Okazaki I., Iweda M., Nojima H.;
CC GAK: a cyclin G associated kinase contains a tensin/auxilin-like

RT domain.";
RL FEBS Lett. 402:73-80(1997).
CC -1- FUNCTION: Associates with cyclin G and CDK5. Seems to act as an
CC auxilin homolog that is involved in the uncoating of clathrin-
CC coated vesicles by Hsc70 in non-neuronal cells. Expression
CC oscillates slightly during the cell cycle, peaking at G1 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Localizes to the perinuclear area and to the
CC trans-Golgi network. Also seen on the plasma membrane, probably at
CC focal adhesions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 TENSIN DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38360; BAA18911.1; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001623; DnaJ N.
DR InterPro: IPR000719; Euk_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF00226; DnaJ; 1.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR SMART: SM00271; DnaJ; 1.
DR SMART: SM00012; PRPC_DSPC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00636; DnaJ_1; FALSE_NEG.
DR PROSITE: PS50076; DnaJ_2; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein; Endoplasmic reticulum; Cell cycle.
FT DOMAIN 40 173
FT ACT_SITE 173 173 PROTEIN KINASE.
FT DOMAIN 405 689 BY SIMILARITY.
FT TENSIN.
FT DOMAIN 1241 1305 J-DOMAIN.
SQ SEQUENCE 1305 AA: 143702 MW: 6D36BD38011C44EE CRC64;

Query Match 49.4%; Score 38; DB 1; Length 1305;
Best Local Similarity 61.5%; Pred. No. 92;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 DLPMSRGVIAN 13
DB 1015 DLPMSRGVIAN 1027

RESULT 12

RA LRP1_CHICK
RT LRP1_CHICK
RL LRP1_CHICK
CC LRP1_CHICK
CC 01-OCT-1996 (Rel. 34, Created)
CC 01-OCT-1996 (Rel. 34, Last annotation update)
CC 16-OCT-2001 (Rel. 40, Last annotation update)
CC Low-density lipoprotein receptor-related protein 1 precursor (LRP)
CC (Alpha-2-macroglobulin receptor) (A2MR).
CC OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
CC NCBI_TaxID=9031;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=White Leghorn; TISSUE=liver, and ovary;
CC MEDLINE=94103212; PubMed=7506255;
CC Nampf J., Stifani S., Bilous P.T., Schneider W.J.;

FT DISULFID 1020 1038 BY SIMILARITY.
 FT DISULFID 1032 1049 BY SIMILARITY.
 FT DISULFID 1060 1073 BY SIMILARITY.
 FT DISULFID 1067 1086 BY SIMILARITY.
 FT DISULFID 1080 1095 BY SIMILARITY.
 FT DISULFID 1102 1116 BY SIMILARITY.
 FT DISULFID 1110 1129 BY SIMILARITY.
 FT DISULFID 1123 1138 BY SIMILARITY.
 FT DISULFID 1143 1157 BY SIMILARITY.
 FT DISULFID 1150 1170 BY SIMILARITY.
 FT DISULFID 1164 1180 BY SIMILARITY.
 FT DISULFID 1183 1194 BY SIMILARITY.
 FT DISULFID 1190 1204 BY SIMILARITY.
 FT DISULFID 1206 1219 BY SIMILARITY.
 FT DISULFID 1225 1235 BY SIMILARITY.
 FT DISULFID 1231 1244 BY SIMILARITY.
 FT DISULFID 1246 1259 BY SIMILARITY.
 FT DISULFID 1538 1551 BY SIMILARITY.
 FT DISULFID 1547 1561 BY SIMILARITY.
 FT DISULFID 1563 1576 BY SIMILARITY.
 FT DISULFID 1846 1857 BY SIMILARITY.
 FT DISULFID 1853 1867 BY SIMILARITY.
 FT DISULFID 1869 1882 BY SIMILARITY.
 FT DISULFID 2155 2166 BY SIMILARITY.
 FT DISULFID 2162 2176 BY SIMILARITY.
 FT DISULFID 2178 2190 BY SIMILARITY.
 FT DISULFID 2476 2487 BY SIMILARITY.
 FT DISULFID 2483 2497 BY SIMILARITY.
 FT DISULFID 2499 2511 BY SIMILARITY.
 FT DISULFID 2518 2531 BY SIMILARITY.
 FT DISULFID 2526 2544 BY SIMILARITY.
 FT DISULFID 2538 2555 BY SIMILARITY.
 FT DISULFID 2560 2572 BY SIMILARITY.
 FT DISULFID 2567 2585 BY SIMILARITY.
 FT DISULFID 2579 2594 BY SIMILARITY.
 FT DISULFID 2599 2611 BY SIMILARITY.
 FT DISULFID 2606 2624 BY SIMILARITY.

Query Match 49.4%; Score 38; DB 1; Length 4543;
 Best Local Similarity 54.5%; Pred. No. 3.6e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLPMSRGVIA 11
 Db 2453 DLPDPMGITA 2463

RESULT 13
 LRP1_HUMAN STANDARD; PRT; 4544 AA.
 AC Q07954;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Low-density lipoprotein receptor-related protein 1 precursor (LRP)
 (Alpha-2-macroglobulin receptor) (A2MR) (Apolipoprotein E receptor)
 (APOE) (C9S1).
 LRP1 OR A2MR.
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 (1)
 SEQUENCE FROM N.A.
 RX TISSUE=LIVER;
 MEDLINE=89210795; PubMed=326596;
 RA Herz J., Hamann U., Rognes S., Myklebost O., Gausepohl H.,
 Stanley K.K.;
 "Surface location and high affinity for calcium of a 500-kd liver
 membrane protein closely related to the LDL-receptor suggest a
 physiological role as lipoprotein receptor.";
 EMBO J. 7:4119-4127(1988).
 (2)

RP SEQUENCE FROM N.A.
 RX MEDLINE=95203893; PubMed=7534747;
 RA Van Leuven F., Stas L., Hilliker C., Lorent K., Unans L., Serneels L.,
 Overbergh L., Torrekens S., Moechars D., De Strooper B.,
 Van den Berghe H.;
 "Structure of the gene (LRP1) coding for the human alpha
 2-macroglobulin receptor lipoprotein receptor-related protein.";
 Genomics 24:78-89(1994).
 (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99000832; PubMed=9782078;
 RA Van Leuven F., Stas L., Thiry E., Nelissen B., Miyake Y.;
 "Strategy to sequence the 89 exons of the human LRP1 gene coding for
 the lipoprotein receptor related protein: identification of one
 expressed mutation among 48 polymorphisms.";
 Genomics 52:138-144(1998).
 (4)
 RP PROCESSING.
 RX MEDLINE=90269210; PubMed=2112085;
 RA Herz J., Kowal R.C., Goldstein J.L., Brown M.S.;
 "Proteolytic processing of the 600 kd low density lipoprotein
 receptor-related protein (LRP) occurs in a trans-Golgi
 compartment.";
 EMBO J. 9:1769-1776(1990).
 (5)
 RP FUNCTION.
 RX MEDLINE=91092405; PubMed=1702392;
 RA Kristensen T., Moestrup S.K., Gilleman J., Bendtsen L., Sand O.,
 Sottrop-Jensen L.;
 "Evidence that the newly cloned low-density-lipoprotein receptor
 related protein (LRP) is the alpha 2-macroglobulin receptor.";
 FEBS Lett. 276:151-155(1990).
 (6)
 RP STRUCTURE BY NMR OF 1059-1100.
 RX MEDLINE=99253972; PubMed=10318830;
 RA Huang W., Dolmer K., Gettins P.G.W.;
 "NMR solution structure of complement-like repeat C88 from the low
 density lipoprotein receptor-related protein.";
 J. Biol. Chem. 274:14130-14136(1999).
 CC -1- FUNCTION: INVOLVED IN THE PLASMA CLEARANCE OF CHYLOMICRON REMNANTS
 AND ACTIVATED ALPHA 2-MACROGLOBULIN, AS WELL AS THE LOCAL
 METABOLISM OF COMPLEXES BETWEEN PLASMINOGEN ACTIVATORS AND THEIR
 ENDOGENOUS INHIBITORS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN LIVER, BRAIN AND LUNG.
 CC -1- PTM: CLEAVED INTO A 85 KDA MEMBRANE-SPANNING SUBUNIT (LRP-85) AND
 A 515 KDA LARGE EXTRACELLULAR DOMAIN (LRP-515) THAT REMAINS NON-
 COVALENTLY ASSOCIATED.
 CC -1- SIMILARITY: CONTAINS 22 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 31 LDL-RECEPTOR CLASS A DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)
 CC or send an email to license@isb-sib.ch).
 DR EMBL: X13916; CA32112.1;
 DR EMBL: AF058427; AAC64265.1;
 DR PDB: 1CR8; 06-JAN-99.
 DR Genew; HGNC:6692; LRP1.
 DR MIM: 107770;
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-like.
 DR InterPro: IPR002172; LDL_receptl.A.
 DR InterPro: IPR000033; ldl_receptor_rep.
 DR Pfam: PF00008; EGF; 16.
 DR Pfam: PF00057; ldl_recept_a; 31.
 DR Pfam: PF00058; ldl_recept_b; 33.
 DR PRINTS: PR00261; LDLRECEPTOR.


```

DR EMBL: AL356332; CAB92055.1; -.
DR HSSP: P11746; 1NMN.
DR TRANSFAC: T03284; -.
DR InterPro: IPR002487; TF_Kbox.
DR InterPro: IPR002100; TF_MADSbox.
DR Pfam: PF00319; SRF-TF; 1.
DR Pfam: PF01486; K-box; 1.
DR PRINTS: PR00404; MADSDOMAIN.
DR SMART: SM00432; MADS; 1.
DR PROSITE: PS00350; MADS_BOX_1; FALSE_NEG.
DR PROSITE: PS50066; MADS_BOX_2; 1.
DR Transcription regulation; Repressor; DNA-binding; Flowering;
KW Developmental protein; Nuclear protein.
FT DOMAIN 3 57
FT DOMAIN 68 166 MADS.
SQ SEQUENCE 196 AA; 21865 MW; 656085A2B157FF42 CRC64;

Query Match 48.1%; Score 37; DB 1; Length 196;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0

OY 1 DLPMSRGVIANSL 14
Db 175 EMBSPAGQISDNL 188

RESULT IS
GAK_MOUSE
ID GAK_MOUSE STANDARD; PRT; 499 AA.
AC 099KRY4.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CYCLIN G-associated kinase (EC 2.7.1.1-) (Fragment).
GN GAK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Associates with cyclin G and CDKs. Seems to act as an
CC auklin homolog that is involved in the uncoupling of clathrin-
CC oscillates slightly during the cell cycle, peaking at G1 (by
CC similarity).
CC -!- SUBCELLULAR LOCATION: Localizes to the perinuclear area and to the
CC trans-Golgi network. Also seen on the plasma membrane, probably at
CC focal adhesions (by similarity).
CC -!- SIMILARITY: BELONGS TO THE SPR/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: BC003958; AAH03958.1; -.
DR InterPro: IPR001623; DnaJ_N.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00226; DnaJ; 1.
DR SMART: SM00271; DnaJ; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; PARTIAL.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; PARTIAL.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; PARTIAL.
DR PROSITE: PS00636; DnaJ_1; FALSE_NEG.

```

DR PROSITE; PS50076; DNAJ_2; 1.
 KW transferase; Serine/threonine-protein kinase; ATP-binding;
 KM Nuclear protein; Endoplasmic reticulum; Cell cycle.
 FT NON TER 1
 FT DOMAIN 435 499 J-DOMAIN
 SO SEQUENCE 499 AA; 53175 MW; 1207D33A92141B96 CRC64;

Query Match 48.1%; Score 37; DB 1; Length 499;
 Best Local Similarity 61.5%; Pred. No. 48;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 DLPMSRGVIAASN 13
 DB 209 DLPAPSKVIASS 221

AC 15
 AC 093JY3; PSEAG STANDARD; PRT; 500 AA.
 AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glucosylglycerol-phosphate synthase (EC 2.4.1.213) (Glucosyl-glycerol-
 phosphate synthase) (GG-phosphate synthase) (GGPS).
 GN GGPS.
 OS Pseudomonas anguilliseptica.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=53406;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hagemann M., Steindruch R.;
 RT "Isolation and characterization of a
 RT glucosylglycerol-phosphate synthase from heterotrophic bacteria. In
 RT preparation."
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in salt tolerance by producing GG-phosphate
 CC from ADP-glucose and glycerol-3-phosphate (G3P), an intermediate
 CC in the synthesis of the osmolyte glucosylglycerol (GG).
 CC -1- CATALYTIC ACTIVITY: ADP-glucose + sn-glycerol 3-phosphate -> 2-
 CC (beta-D-glucosyl)-sn-glycerol 3-phosphate + ADP.
 CC -1- PATHWAY: Glucosylglycerol biosynthesis; first step.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 20.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; AJ18784; CAC50077.1; -
 DR InterPro; IPR001830; GT_20.
 DR Pfam; PF00982; Glyco_transf_20; 1.
 KW Transferase; Glycosyltransferase.
 SO SEQUENCE 500 AA; 56849 MW; D0B85214278BD7DA CRC64;

Query Match 48.1%; Score 37; DB 1; Length 500;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 PMSRGVIAASNLXF 16
 DB 23 PTPMGITPTLSLF 36

RESULT 17
 ID DNAK_BACME STANDARD; PRT; 604 AA.
 AC P05646;

DT 01-NOV-1988 (Rel. 09, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
 DE protein) (HSP70).
 GN DNAK.
 OS Bacillus megaterium.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1404;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=87231083; PubMed=3035506;
 RA Sussman M.D., Seelow P.;
 RT "Nucleotide sequence of a Bacillus megaterium gene homologous to the
 RT dnaK gene of Escherichia coli."

RL Nucleic Acids Res. 15:3923-3923(1987).
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -1- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; Y00154; CAA68348.1; -
 DR HSP: P04475; IDG4.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR PRODOM; PD000089; HSP70; 1.
 DR PROSITE; PS00297; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
 FT INIT_MET 0
 FT MOD_RES 172 172
 FT MOD_RES 172 172
 SO SEQUENCE 604 AA; 65119 MW; 29DF856DC5FAC14 CRC64;

Query Match 48.1%; Score 37; DB 1; Length 604;
 Best Local Similarity 43.8%; Pred. No. 60;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 DLPMSRGVIAASNLXF 16
 DB 430 DLPAPGVQPIEVSF 445

AC 09KMS7; STANDARD; PRT; 606 AA.
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
 DE protein) (HSP70).
 GN DNAK.
 OS Bacillus thermoglucosidasius.
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.

OX NCBI_TaxID=1426;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=KPI006;
 RA Matanabe K., Iwashiro T., Suzuki Y.;
 RT Features of dnaK operon genes of the obligate thermophile Bacillus
 RT thermoglucosidasius KPI006.
 RT Antonie Van Leeuwenhoek 77:241-250(2000).
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -1- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB017035; BAB03215.1; -
CC HSP70; P04475; IDG4.
CC InterPro: IPR001023; Hsp70.
CC Pfam: PF00012; HSP70; 1.
CC PRINTS: PR00301; HEATSHOCK70.
CC ProDom: PD000089; HSP70.1.
CC PROSITE: PS00297; HSP70_1; 1.
CC PROSITE: PS00329; HSP70_2; 1.
CC PROSITE: PS01036; HSP70_3; 1.
CC Chapterone: ATP-binding; Heat shock; Phosphorylation.
CC INIT_MET 0
CC MOD_RES 168 168
CC SEQUENCE 606 AA; 65951 MW; 7955FD3D845F56 CRC64;
CC
CC Query Match
CC Best Local Similarity 48.1%; Score 37; DB 1; Length 606;
CC Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 1 DLPMSRGVYASIMXLF 16
CC Db 426 DIPAPRGVQIEVTF 441
CC
CC RESULT 19
CC DNK_LACIA STANDARD: PRT; 607 AA.
AC P42368; O9CGY8; 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chapterone protein dnak (Heat shock protein 70) (Heat shock 70 kDa
DE protein) (HSP70).
GN DNK OR IL0954.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis), and
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1356; 1359;
RN [1]
RP SPECIES FROM N.A.
RC SPECIES=L.1.cremoris; STRAIN=MG1363;
RX MEDLINE=94172317; PubMed=8126443;
RA Baton T.J., Shearman C.A., Gasson M.J.;
RT "Cloning and sequence analysis of the dnak gene region of Lactococcus
RT lactis subsp. lactis";
RL J. Gen. Microbiol. 139:3253-3263(1993).
RN [12]
RP SEQUENCE FROM N.A.
RC SPECIES=L.1.lactis; STRAIN=LM0230;
RX MEDLINE=94237496; PubMed=8181763;
RA Barril M.J.S., Kim S.G., Batt C.A.;
RT "Cloning and sequencing of the Lactococcus lactis subsp. lactis dnak
RT gene using a PCR-based approach";
RL Gene 142:91-96(1994).
RN [13]
RP SEQUENCE FROM N.A.
RC SPECIES=L.1.lactis; STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winkler P., Mauger S., Tallon O., Malarme K.,
RT Weissbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).

CC -1- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X76642; CA54089.1; -
CC EMBL: X75428; CA53179.1; -
CC HSP70; P04475; IDG4.
CC InterPro: IPR001023; Hsp70.
CC Pfam: PF00012; HSP70; 1.
CC PRINTS: PR00301; HEATSHOCK70.
CC ProDom: PD000089; HSP70.1.
CC PROSITE: PS00297; HSP70_1; 1.
CC PROSITE: PS00329; HSP70_2; 1.
CC PROSITE: PS01036; HSP70_3; 1.
CC Chapterone: ATP-binding; Heat shock; Phosphorylation.
CC MOD_RES 173 173
CC SEQUENCE 607 AA; 64948 MW; 567FA7160313310E CRC64;
CC
CC Query Match
CC Best Local Similarity 48.1%; Score 37; DB 1; Length 607;
CC Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 1 DLPMSRGVYASIMXLF 16
CC Db 431 DIPAPRGVQIEVTF 446
CC
CC RESULT 20
CC DNK_STRPN STANDARD: PRT; 607 AA.
AC P55829; O66035; 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chapterone protein dnak (Heat shock protein 70) (Heat shock 70 kDa
DE protein) (HSP70).
GN DNK OR SP0517.
OS Streptococcus pneumoniae.
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. pneumoniae; STRAIN=ATCC 6305; 15-JUL-1998 (Rel. 36, Created)
RX MEDLINE=98231633; PubMed=9570114;
RA Kim S.-W., Choi I.-H., Kim S.-N., Kim Y.-H., Pyo S.-N., Rhee D.-K.;
RT "Molecular cloning, expression, and characterization of dnak in
RT Streptococcus pneumoniae";
RL FEMS Microbiol. Lett. 161:217-224(1998).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=TTGR4;


```

RX MEDLINE=21357209; PubMed=11463916;
RA Tetteil H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angluoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RA "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; U72720; AAB39221.1; -.
CC EMBL; U84387; AAC15892.1; -.
CC EMBL; AE007363; AAK74675.1; -.
CC HSP; P04475; IDG4.
CC TIGR; SP0517; -.
CC
CC InterPro: IPR001023; HSP70.
CC Pfam; PF00012; HSP70; 1.
CC PRINTS; PR00301; HEATSHOCK70.
CC ProDom; PD000089; HSP70_1.
CC PROSITE; PS00297; HSP70_1; 1.
CC PROSITE; PS00329; HSP70_2; 1.
CC PROSITE; PS01036; HSP70_3; 1.
CC
CC K1W Chaperone; ATP-binding; Heat shock; Phosphorylation;
CC Complete proteome.
CC
CC MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC FT CONFLICT 493 493 A -> P (IN REF. 2).
CC FT CONFLICT 495 495 A -> S (IN REF. 1).
CC FT CONFLICT 530 530 A -> P (IN REF. 2).
CC FT CONFLICT 556 556 T -> A (IN REF. 1 AND 2).
CC FT CONFLICT 601 602 DG -> E (IN REF. 2).
CC SQ SEQUENCE 607 AA; 64842 MM; 11D626F1B37D0760 CRC64;

```

Search completed: January 2, 2003, 12:03:45
Job time : 35 secs

```

Query Match 48.18; Score 37; DB 1; Length 607;
Best Local Similarity 37.38; Pred. No. 60;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
DB 431 DIPNAPRGIPQIEVTF 446
OY 1 DLPSPRGVINSNLF 16
1:1:111:1:1

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 12:03:04 ; Search time 100 Seconds

(without alignments)
32.968 Million cell updates/sec

Title: US-09-824-053-3

Sequence: 1 DLPMSRGVYASNLXF 16

Gaping table: GAPOP 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 671580

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHA:21:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MNC:*
8: SP_ORNITHINE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Match length DB ID Description

75 97.4 546 10 P93762 Chondrus cr
44 57.1 1096 11 Q9D504 mus musculu
44 57.1 1207 11 Q9D5V1 mus musculu
43 55.8 184 16 Q8U8M0 agrobacteri
5 55.8 631 5 O17051 ascaris lum
41 53.2 416 16 O8XR29 escherichia
7 41 1051 3 P97961 cyllindrotri
8 40 51.9 229 13 Q90235 ambystoma m
9 40 51.9 407 17 Q9HPM4 halobacteri
10 40 51.9 989 5 Q9N9Z1 sulfolobus
11 40 51.9 1193 4 Q92580 oryza sativ
12 40 51.9 64 10 Q943R9 oryza sativ
13 39 50.6 112 10 Q943K8 oryza sativ
14 39 50.6 113 13 Q91153 notophthalm
15 39 50.6 113 13 Q91153 notophthalm
16 39 50.6 150 10 O8S724 oryza sativ

17 39 50.6 280 13 Q91152 notophthalm
18 39 50.6 442 16 Q981L6 rhizobium l
19 39 50.6 468 4 Q9NMB8 homo sapien
20 39 50.6 468 4 Q9BX59 homo sapien
21 39 50.6 541 5 Q9Y1L1 locusta mig
22 39 50.6 1143 10 Q948Y6 volvox cart
23 38 49.4 51 4 Q9BXH2 homo sapien
24 38 49.4 51 11 Q9NML3 mus musculu
25 38 49.4 211 2 Q9VW46 Drevibacter
26 38 49.4 328 16 Q92XH5 rhizobium m
27 38 49.4 476 5 Q9N8C7 trypansom
28 38 49.4 479 16 Q949746 mycobacteri
29 38 49.4 534 17 Q97CT0 thermoplas
30 38 49.4 618 2 Q93R27 tetragenoc
31 38 49.4 622 10 Q94IX1 oryza sativ
32 38 49.4 670 16 Q9RXQ7 delinococcus
33 38 49.4 684 17 Q97UR0 sulfolobus
34 38 49.4 769 13 Q9BWS4 homo sapien
35 38 49.4 769 13 Q9PVX8 xenopus lae
36 38 49.4 770 4 Q94916 homo sapien
37 38 49.4 890 10 Q9RY23 arabidopsis
38 38 49.4 3223 2 Q9RPH1 escherichia
39 38 49.4 3223 2 Q9RPH1 escherichia
40 38 49.4 3223 2 Q9RPH1 escherichia
41 38 49.4 3223 2 Q9RPH1 escherichia
42 38 49.4 4545 11 Q920Y4 mus musculu
43 38 49.4 4545 11 Q920Y4 mus musculu
44 38 49.4 4545 11 Q912X7 mus musculu
45 37 48.1 245 16 Q820S3 anabaena sp

ALIGNMENTS

RESULT 1

P93762 PRELIMINARY: PRT; 546 AA.
AC P93762; 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DI 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hexose oxidase.
OS Chondrus crispus (Carrageen).
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartineaceae;
OC Chondrus.
OX NCBI_TaxID-2769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97269074; PubMed-9111074;
RA Hansen O.C., Stougaard P.;
RT "Hexose oxidase from the red alga Chondrus crispus. Purification,
RT molecular cloning, and expression in Pichia pastoris.";
RL J. Biol. Chem. 272:11581-11587(1997).
DR EMBL: U089770; AAA09376.1;
DR InterPro: IPR001575; Oxid_FAD_bind.
DR Pfam: PF01565; FAD_binding_4; 1.
SQ SEQUENCE 546 AA; 61899 MW; 8DDEDA46D53C870B1 CRC64;

Query Match 97.4% Score 75; DB 10; Length 546;
Best Local Similarity 93.8% Pred. No. 4.5e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DLPMSRGVYASNLXF 16
DB 219 DLPMSRGVYASNLXF 234

RESULT 2
Q9D504 PRELIMINARY: PRT; 1096 AA.
AC Q9D504; 01-JUN-2001 (TREMBLrel. 17, Created)
DI 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE 4921517D2IRIK protein.
 GN 4921517D2IRIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK014926; BAB29623.1; -
 DR MGD: MGI:1914972; 4921517D2IRIK.
 DR InterPro: IPR004000; Actin_like.
 DR Pfam: PF00022; actin_1.
 DR SMART: SM00268; ACTIN_1.
 SQ SEQUENCE 1096 AA; 119041 MW; EB5E1F51D0A1C983 CRC64;
 Query Match 57.1%; Score 44; DB 11; Length 1096;
 Best Local Similarity 81.8%; Pred. No. 35;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LPMSPRGVAS 12
 DB 289 LPSPRGVIVS 299
 RESULT 3
 ID 09D5V1 PRELIMINARY; PRT; 1207 AA.
 AC 09D5V1;
 RT 01-JUN-2001 (TREMblrel. 17, Created)
 RT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 RT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 RT 4921517D2IRIK protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK014908; BAB29615.1; -
 DR MGD: MGI:1914972; 4921517D2IRIK.
 DR InterPro: IPR004000; Actin_like.
 DR Pfam: PF00022; actin_1.
 DR SMART: SM00268; ACTIN_1.
 SQ SEQUENCE 1207 AA; 130803 MW; 2CA22CD754C0F5D8 CRC64;
 Query Match 57.1%; Score 44; DB 11; Length 1207;
 Best Local Similarity 81.8%; Pred. No. 36;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LPMSPRGVAS 12
 DB 400 LPSPRGVIVS 410
 RESULT 4
 ID 08U8M0 PRELIMINARY; PRT; 184 AA.
 AC 08U8M0;
 RT 01-JUN-2002 (TREMblrel. 21, Created)
 RT 01-JUN-2002 (TREMblrel. 21, last sequence update)
 RT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 RT Sarcosine oxidase gamma subunit.
 DE Sarcosine oxidase gamma subunit.
 GN SOXG OR ATU4070 OR AGR.L.1572.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Moors D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D., St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Doan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RT C58.";
 RT Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Quicilo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaubin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cleto C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RT Science 294:2323-2328(2001).
 RL EMBL: AE009338; AA144871.1; -
 DR EMBL: AE008277; AAK69355.1; -
 DR Complete proteome.
 KW SEQUENCE 184 AA; 19159 MW; 73D49565A448580B CRC64;

```

Query Match      55.8%   Score 43; DB 16; Length 144;
Best Local Similarity 53.3%   Pred. No. 8.3;
Matches      8; Conservative      2; Mismatches      5; Indels      0; Gaps      0.
OY      2      LPMSPRCVIAASLXLP      16
      ||||:|:|:|
DB      54      LPTSPKSGVTSGLRT      68

```

```

RESULT 5
ID 017051 PRELIMINARY: PRT: 631 AA.
AC 017051:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
Gag protein.
Ascaris lumbricoideae (common roundworm).
Eukaryotes; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea;
Ascarididae; Ascaris.
NCBI_Taxid=6252;
OC
[1]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=95047480; PubMed=7525414;
RA Felder H., Herczeg A., de Chastonay Y., Aeby P., Tobler H., Muller F.;
RT "rs", a protretrovirus-like element from the parasitic nematode Ascaris
RL lumbricoideae."
RN Gene 149:219-225(1994).
RN [2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=87133498; PubMed=3816762;
RA Aeby P., Spicher A., de Chastonay Y., Mueller F., Tobler H.;
RT "Structure and genomic organization of protretrovirus-like elements
RT partially eliminated from the somatic genome of Ascaris
RL lumbricoideae."
RN EMBO J. 5:3353-3360(1986).
RN [3]
RN
RP SEQUENCE FROM N.A.
RA Helinz F.H.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z29712; CAA82797.1; -.
DR MEROPS; A16.001; -.
DR InterPro: IPR005312; Peptidase_A16.
DR InterPro: IPR001878; ZnF_CCHC.
DR Pfam; PF03564; Peptidase_A16; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; ZnF_C2HC; 2.
DR SEQUENCE 631 AA; 70684 MW; DBD9260982705F60 CRC64;
Query Match 55.8%; Score 43; DB 5; Length 631;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0
OY 1 DLPMSRGVYASNL 14
DB 432 DQPLSPDEEYIANM 445
RESULT 6
ID 08XB29 PRELIMINARY: PRT: 416 AA.
AC 08XB29:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN YIDR OR Z5185 OR ECS4629.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=83334;
RN
RP SEQUENCE FROM N.A.

```

RA	STRAIN-0157:H7 / EDL933 / ATCC 700927;
RX	MEDLINE-21074935; PubMed-11206551;
RA	Perna N.T., Plunkett G., Ili, Burland V., Mau B., Glasner J.D.,
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA	Großbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamousis K.,
RA	Apotheca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA	Weich R.A., Blattner F.R.;
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"
RL	Nature 409:529-533(2001).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN-O157:H7 / RIMD 050952:
RX	MEDLINE-21135231; PubMed-11238796;
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA	Iida T., Takami H., Honda T., Sasakiwa C., Ogawara N., Yasunaga T.,
RA	Kuhara S., Shibata T., Hattori M., Shinagawa H.;
RT	"Complete genome sequence of enterohemorrhagic Escherichia coli
RT	O157:H7 and genomic comparison with a laboratory strain K-12."
RL	DNA Res. 8:11-22(2001).
DR	EMBL; AE005600; AAC58891.1; "
DR	EMBL; AP002566; BAB38052.1; "-
KW	Complete proteome.
SQ	SEQUENCE 416 AA; 46314 MW; 37DI66B52A88669 CRC64;
<hr/>	
Query Match	53.2%; Score 41; DB 16; Length 416;
Best Local Similarity	46.7%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;	
OY	2 LPMSRGVYASNLNF 16 ::
DB	272 LPAPRGIVORRLTF 286
<hr/>	
RESULT 7	
ID	P97961 PRELIMINARY; PRT; 1051 AA.
AC	P97961;
DT	01-MAY-1997 (TREMBLrel. 03, Created)
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	Peptidase synthetase (Fragment).
OS	Cylindrocapsa oligosperma.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC	Sordariales; Chaetosphaeriales; Chaetosphaeriaceae;
OC	Cylindrocapsa
OX	NCBI_TaxID=72418;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-NRRL 18230;
RA	Bernhard F.;
RT	"Identification of genes encoding for peptide synthetases from Gram-
RT	negative bacteria and filamentous fungi,"
RL	Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
DR	EMBL; X96559; CAA65395.1; "-
DR	HSSP; P14687; IAMU.
DR	InterPro; IPRO00873; AMP-bind.
DR	InterPro; IPRO01242; Condensatn.
DR	InterPro; IPRO03880; Praptnec.attach.
DR	pFam; PF00501; AMP-binding_2.
DR	pFam; PF00668; Condensation_1.
DR	pFam; PF00550; pp-binding_1.
DR	PRINTS; PR00154; AMPBINDING.
DR	PROSITE; PS0075; ACP_DOMAIN_1.
DR	PROSITE; PS00455; AMP_BINDING_1.
DR	PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
FT	NON_TER 1
FT	NON_TER 1051 1051
SQ	SEQUENCE 1051 AA; 116526 MW; 7C8926576B2F0619 CRC64;
<hr/>	
Query Match	53.2%; Score 41; DB 3; Length 1051;
Best Local Similarity	60.0%; Pred. No. 1.le+02;

RESULT 7			
ID	PRELIMINARY:	PRT:	1051 AA.
DT	01-MAY-1997 (TREMBLrel. 03, Created)		
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Peptidase synthetase (Fragment).		
OS	Cylindrocapsa oligosperma.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariaceae; Chaetosphaeriaceae; Chaetosphaeriaceae; Chaetosphaerium.		
OC	Cylindrocapsa.		
OX	NCBI_Taxid=72418;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NRRL 18230;		
RA	Bernhard F.;		
RT	"Identification of genes encoding for peptide synthetases from Gram-negative bacteria and filamentous fungi.";		
RL	Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.		
EMBL	X96559: CAA65395.1; -.		
DR	HSSP; P14687; 1AMU.		
DR	InterPro; IPR000873; AMP-bind.		
DR	InterPro; IPR001242; Condensatn.		
DR	InterPro; IPR003880; Pantne_attach.		
DR	Pfam; PF00501; AMP-binding; 2.		
DR	Pfam; PF00668; Condensation; 1.		
DR	Pfam; PF00550; pp-binding; 1.		
DR	PRINTS; PR00154; AMPBINDING.		
DR	PROSITE; PS50075; ACP_DOMAIN; 1.		
DR	PROSITE; PS00455; AMP_BINDING; 1.		
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.		
FT	NON_TER	1	
FT	NON_TER	1051	
FT	SEQUENCE	1051 AA;	116526 MW; 7C8926576B2F0619 CRC64;
Query Match	53.2%;	Score 41;	DB 3; Length 1051;
Best Local Similarity	60.0%;	Pred. No. 1.1e+02;	

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 LPMSPRGVIANLXF 16
|||||
Db 936 LPMTPRVAHLSNLCF 950

RESULT 8

Q90235 PRELIMINARY; PRT; 229 AA.
AC 090235;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mx1 protein.
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
[1]

SEQUENCE FROM N.A.

TISSUE=REGENERATING BLASTEMA;
Koshiba K., Yamamoto H., Tamura K., Ide H., Kuroiwa A.;
"Expression of Mx genes in regenerating and developing limbs of axolotl.";
Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC EMBL: DB2577; BAA11574.1; -
DR HSSP: P14653; 1B72.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox.1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR PRINTS: PR00024; HOMEBOX.
DR Prodom: PD000010; Homeobox.1.
DR SMART: SM00389; HOX.1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 229 AA; 25550 MW; 81BAD98CFEER015 CRC64;

Query Match 51.9%; Score 40; DB 13; Length 229;
Best Local Similarity 40.0%; Pred. No. 36;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPMSPRGVIANLXF 16
|||||
Db 209 MPMPGGLAHMGY 223

RESULT 9

Q9HPM4 PRELIMINARY; PRT; 407 AA.
Q9HPM4;
01-MAR-2001 (TREMBLrel. 16, Created)
01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Vng1564h.
GN VNG1564h.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacterla; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
[1]

SEQUENCE FROM N.A.

RX MEDLINE-20504483; PubMed-11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Sirogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angelvine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebbhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005068; AAC19843.1; -
KW Complete proteome.
SQ SEQUENCE 407 AA; 42994 MW; 2ED7DBB35C2C1380 CRC64;

Query Match 51.9%; Score 40; DB 17; Length 407;
Best Local Similarity 46.2%; Pred. No. 65;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPMSPRGVIANL 13
|:|||||:
Db 223 DVALLPRGVAAAS 235

RESULT 10

Q974B5 PRELIMINARY; PRT; 682 AA.
AC 0974B5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative carbon-monoxide dehydrogenase large subunit.
GN S70739.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
[1]

SEQUENCE FROM N.A.

RP STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000983; BAB65745.1; -
DR InterPro: IPR000674; Aldxan_dh.C.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF01315; Ald_xan_dh.C.1.
DR Pfam: PF02738; Ald_xan_dh.C2.1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 682 AA; 74690 MW; 863B40E9B318AA6F CRC64;

Query Match 51.9%; Score 40; DB 17; Length 682;
Best Local Similarity 70.0%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PMSPRGVIAN 12
|:|||||:
Db 172 PMEPKGVIAN 181

RESULT 11

Q9N9Z1 PRELIMINARY; PRT; 989 AA.
AC 09N9Z1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Endonuclease/reverse transcriptase.
GN PILGERPOL OR POL.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea: Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-PIIIGER NON-LTR RETROTRANSPOSON;
 RA Tuschel T., Elbashir S., Lendeckel W.;
 RT "Drosophila melanogaster non-LTR retrotransposon pIIGER";
 RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A0278684; CAB99192.1;
 DR FlyBase: FBgn0041726; pIIGER-pol.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF03372; Exo_endo_phos. 1.
 DR Pfam: PF00078; rvc. 1.
 KM RNA-directed DNA polymerase.
 SQ SEQUENCE 989 AA; 110315 MW; 5957679C395D79A6 CRC64;
 Query Match 51.9%; Score 40; DB 5; Length 989;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 OY 3 PMSPRGVAS 12
 DB 947 PLSPRGIVAA 956
 RESULT 12
 ID 092580 PRELIMINARY; PRT; 1193 AA.
 AC 092580;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE KIA00268 protein (Fragment).
 GN KIA00268.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BONE MARROW;
 RA MEDLINE-97191544; PubMed-9039502;
 RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
 RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RT Prediction of the coding sequences of unidentified human genes. VI.
 RT the coding sequences of 80 new genes (KIA0201-KIA0280) deduced by
 RT analysis of cDNA clones from cell line KG-1 and brain.*;
 RT DNA Res. 3:321-329(1996).
 RT EMBL: D87742; BAA13448.1; .
 SQ SEQUENCE 1193 AA; 133704 MW; 70FB2542F9998038 CRC64;
 NON_TER 1
 Query Match 51.9%; Score 40; DB 4; Length 1193;
 Best Local Similarity 43.8%; Pred. No. 2e+02;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 OY 1 DLPMSPRGVASLXAF 16
 DB 1113 DLPMPRGFLGHPAF 1128
 RESULT 13
 ID 094389 PRELIMINARY; PRT; 64 AA.
 AC 094389;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE OSJNB0083M16.17 protein.
 GN OSJNB0083M16.17.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
 RT clone:OSJNB0083M16.*";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003214; BAB64617.1;
 SQ SEQUENCE 64 AA; 7195 MW; F0FD27343FC0C920 CRC64;
 Query Match 50.6%; Score 39; DB 10; Length 64;
 Best Local Similarity 87.5%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 LPMSPRGV 9
 DB 5 LPMSPRGV 12
 RESULT 14
 ID 0943K8 PRELIMINARY; PRT; 112 AA.
 AC 0943K8;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P0031D11.27 protein.
 GN P0031D11.27.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0031D11.*";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003231; BAB67893.1; .
 SQ SEQUENCE 112 AA; 12233 MW; B750B7CA83A448D6 CRC64;
 Query Match 50.6%; Score 39; DB 10; Length 112;
 Best Local Similarity 87.5%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 LPMSPRGV 9
 DB 5 LPMSPRGV 12
 RESULT 15
 ID 091153 PRELIMINARY; PRT; 113 AA.
 AC 091153;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Transcription factor (Fragment).
 GN MSX-1.
 OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
 OC Notoththalmus.
 OX NCBI_TaxID=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-MID-BUD OF REGENERATING LIMB;
 RT MEDLINE-95218226; PubMed-7703517;

RA Simon H.G., Nelson C., Goff D., Laufer E., Morgan B., Tabin C.;
 RT "Differential expression of myogenic regulatory genes and Msx-1 during
 RT dedifferentiation and redifferentiation of regenerating amphibian
 RT limbs";
 RL Dev. Dyn. 202;1-12(1995).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: X82837; CA58044.1; -.
 DR HSSP: P22808; INK3.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox.1.
 DR PRINTS: PR00024; HOMEBOX.1.
 DR ProDom: PD000010; Homeobox.1.
 DR SMART: SM00389; HOX.1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 KW NON_TER
 FT
 SQ SEQUENCE 113 AA; 12732 MW; 13E442F4858D982 CRC64;

Query Match 50.6%; Score 39; DB 13; Length 113;
 Best Local Similarity 40.0%; Pred. No. 26;
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

2 LPMSPRGVIAASNLXF 16
 :||||| : : : : :
 Db 93 MPMSPMGLYAAHVGY 107

RESULT 16

Q8S724 PRELIMINARY; PRT; 150 AA.

AC Q8S724;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 16.6 kDa protein.
 GN OSUNB0091N21.24.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_Taxid=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Gasberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tsitrin T., Riggs F., Hsiao J., Zisman V., Blunt S., Pal G.,
 RA VanKen S.E., Utterback T.R., Feldblum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSUNB0091N21 genomic sequence.";
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 EMBL: AC091122; AAL2695.1; -.
 Hypothetical protein.
 SEQUENCE 150 AA; 16595 MW; 9752C590F7DE8118 CRC64;

Query Match 50.6%; Score 39; DB 10; Length 150;
 Best Local Similarity 87.5%; Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPMSPRGV 9
 ||:|||||
 Db 5 LPLSPRGV 12

RESULT 17

Q91152 PRELIMINARY; PRT; 280 AA.

AC Q91152;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Msx-1 protein.

GN MSX-1.
 OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
 OC Notoththalmus.
 OX NCBI_Taxid=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=REGENERATING BLASTEMA;
 RX MEDLINE=95249617; PubMed=7732036;
 RA Crews L., Gates P., Brown R., Joliot A., Foley C., Brookes J.P.,
 RA Gann A.A.;
 RT "Expression and activity of the newt Msx-1 gene in relation to limb
 RT regeneration.";
 RL Proc. R. Soc. Lond. B, Biol. Sci. 259:161-171(1995).
 DR EMBL: X82395; CA57791.1; -.
 DR HSSP: P14653; 1B72.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox.1.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD000010; Homeobox.1.
 DR SMART: SM00389; HOX.1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 KW
 FT
 SQ SEQUENCE 280 AA; 30966 MW; B8BBD462003CA87B CRC64;

Query Match 50.6%; Score 39; DB 13; Length 280;
 Best Local Similarity 40.0%; Pred. No. 67;
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPMSPRGVIAASNLXF 16
 :||||| : : : : :
 Db 260 MPMSPMGLYAAHVGY 274

RESULT 18

Q98116 PRELIMINARY; PRT; 442 AA.

AC Q98116;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein mlr2348.
 GN MLR2348.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_Taxid=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Saito S., Asamizu E., Kato T., Sasamoto S.,
 RA Kishida Y., Kiyokawa C., Ishikawa A., Kawashima K., Kimura T.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002999; BAB49500.1; -.
 DR InterPro: IPR001145; Bac.OmpA.
 DR Pfam: PF00691; OmpA; 1.
 DR PRINTS: PR01023; NAFLGOMY.
 DR ProDom: PD000930; Bac.OmpA; 1.
 KW Hypothetical protein; Complete proteome.
 FT
 SQ SEQUENCE 442 AA; 47860 MW; 4853A307B1D71850 CRC64;

Query Match 50.6%; Score 39; DB 16; Length 442;
 Best Local Similarity 60.0%; Pred. No. 11e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 DLPMSPP--RGVIAASN 13


```

Db      205 DAPLSPMKGVIASD 219

RESULT 19
ID      09NMB8      PRELIMINARY;      PRT;      468 AA.
AC      09NMB8;
DT      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT      01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE      Hypothetical 50.2 kDa protein.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=EMBRYO;
RA      Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA      Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA      Watanabe S., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA      Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA      Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA      Yamamoto J., Wakematsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA      Ninomiya K., Iwayanagi T.;
RT      "NEDO human cDNA sequencing project.";
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK001005; BAB91465.1; -
DR      InterPro; IPR003597; IG-cl.
DR      InterPro; IPR003600; IG-like.
DR      InterPro; IPR003006; IG_MHC.
DR      Pfam; PF00047; Ig_2.
DR      SMART; SM00407; IGcl; 1.
DR      SMART; SM00410; IG-like; 1.
DR      PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ      SEQUENCE 468 AA; 50220 MW; 7C2FC6D5744FB101 CRC64;

Query Match      50.6%; Score 39; DB 4; Length 468;
Best Local Similarity 37.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      1 DLPMSPRGVIASMLXF 16
Db      176 NLPLSPQGVTRAVEF 191

ULT 20
X59
09BX59 PRELIMINARY; PRT; 468 AA.
AC      09BX59;
DT      01-JUN-2001 (TREMBlrel. 17, Created)
DT      01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT      01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE      PLACEL007632 protein (Hypothetical 50.2 kDa protein).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=PLACENTA;
RA      Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA      Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA      Watanabe S., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA      Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA      Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA      Yamamoto J., Wakematsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA      Ninomiya K., Iwayanagi T.;
RT      "NEDO human cDNA sequencing project.";
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.

```

```

RC      TISSUE=SKIN;
RA      Strausberg R.;
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AK002056; BAB41077.1; -
DR      EMBL; BC015017; AAH15017.1; -
DR      InterPro; IPR003597; IG-cl.
DR      InterPro; IPR003597; IG-cl.
DR      InterPro; IPR003006; IG_MHC.
DR      Pfam; PF00047; Ig_2.
DR      SMART; SM00409; IG_1.
DR      SMART; SM00407; IGcl; 1.
DR      PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW      Hypothetical protein.
SQ      SEQUENCE 468 AA; 50177 MW; 1826E2C8F3C39841 CRC64;

Query Match      50.6%; Score 39; DB 4; Length 468;
Best Local Similarity 37.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      1 DLPMSPRGVIASMLXF 16
Db      176 NLPLSPQGVTRAVEF 191

```

Search completed: January 2, 2003, 12:07:20
Job time : 103 secs

